The shortest distance in random multi-type intersection graphs

Barbour, A D ; Reinert, G

Abstract: Using an associated branching process as the basis of our approximation, we show that typical inter-point distances in a multi-type random intersection graph have a defective distribution, which is well described by a mixture of translated and scaled Gumbel distributions, the missing mass corresponding to the event that the vertices are not in the same component of the graph. © 2010 Wiley Periodicals, Inc.

DOI: https://doi.org/10.1002/rsa.20351

Originally published at:
Barbour, A D; Reinert, G (2011). The shortest distance in random multi-type intersection graphs. Random Structures Algorithms, 39(2):179-209.
DOI: https://doi.org/10.1002/rsa.20351

The shortest distance in random multi-type intersection graphs

A. D. Barbour* and G. Reinert†
Universität Zürich and University of Oxford

Abstract

Using an associated branching process as the basis of our approximation, we show that typical inter-point distances in a multitype random intersection graph have a defective distribution, which is well described by a mixture of translated and scaled Gumbel distributions, the missing mass corresponding to the event that the vertices are not in the same component of the graph.

Keywords. Intersection graph, shortest path, branching process approximation, Poisson approximation.

1 Introduction

Bipartite graphs have been studied in a variety of applications: directors and companies [19], persons and questions in an intelligence test [17], or genes and gene properties [21], to give just a few examples. Typically, in such applications, the graph induced on the vertices of one of the two parts, with vertices linked if there is a path of length 2 joining them in the bipartite graph, are of primary interest. For instance, the structure of the network linking directors may be of greater interest than the companies involved. Furthermore, in some applications, the remaining part of the bipartite graph, which is responsible for forming the links, may not be known or observable, and it may be of interest to deduce its existence from the properties of the observed part of the structure alone. However, the statistical properties of such bipartite graphs are not well understood, particularly when there are different types of vertices, see [19]. Here, we shall be concerned with the properties of a particular family of such graphs, known as random intersection graphs, and with the statistics of distances between randomly chosen points.

Random intersection graphs are constructed from two sets, the 'vertices' and the 'objects', as follows. Each vertex $v \in V$ is associated with a randomly chosen subset A_v of

*Angewandte Mathematik, Universität Zürich, Winterthurertrasse 190, CH-8057 ZÜRICH; ADB was supported in part by Schweizer Nationalfonds Projekt Nr. 20–117625/1, and carried out part of the work while visiting the Institut Mittag–Leffler, Djursholm, Sweden.

†Department of Statistics, University of Oxford, 1 South Parks Road, OXFORD OX1 3TG, UK; GR was supported in part by EPSRC and BBSRC through OCISB.

a finite set U of objects, and two vertices v and v' are joined in the graph if $A_v \cap A_{v'} \neq \emptyset$. In the simplest case, the Bernoulli model, vertex v is associated with object u independently of all other associations with fixed probability p. Britton *et al.* [6] establish a branching process approximation for the spread of a Reed-Frost epidemic on such a graph. Here, we consider the more flexible model in which there are K distinct types of vertices and J types of objects, and vertex v of type k is associated with object u of type j independently of all other associations with probability p_{kj}. Our model can be viewed as a bipartite Erdős–Rényi mixture graph [8]. In Erdős–Rényi mixture graphs, vertices are coloured, with the probability of two vertices being connected depending only on their colours; edges occur independently.

Random intersection graphs of this kind can also be related to the Rasch [17] models in social science. These are given by taking

$$p_{kj} = \frac{\alpha_k \beta_j}{1 + \alpha_k \beta_j}.$$

For example, one might have $k \sim j$ if person k solves problem j correctly; the α's would then relate to the ability of the person, and the β's to the type of problem. A simplified Rasch model of the form $p_{kj} = \alpha_k \beta_j$ can be viewed as a special case of an exponential random graph model, see Equation (1) in [20].

In the study of random networks, the shortest distance between two randomly chosen vertices is one of the standard summary statistics. In this paper, we approximate its distribution for multitype Bernoulli intersection graphs. Since the networks used in applications are typically finite, we not only provide a limiting approximation, but also give explicit bounds on the difference between the true and limiting distributions. Our main results, summarized in Corollary 6.4, give an approximation described in terms of the distributions of the limiting random variables W of the associated multivariate bipartite branching process, when the process starts with a single individual of one or other of the types. The probability of the two vertices being in the same component is well approximated by the product of the probabilities that neither of the branching processes becomes extinct. On this event, the distance has a distribution close to that of a translation–mixture of scaled Gumbel distributions, with the mixture distribution being explicitly given in terms of those of the limiting random variables W. Alternatively, the approximate distribution can be described as that of (a linear transformation of) the sum of three independent random variables, one a Gumbel, and the others distributed as the logarithm of W, given the appropriate initial types. In a natural asymptotic framework, the error bounds behave like an inverse power of the total number of vertices, whose exponent can be derived from the parameters of the bipartite graph: the probabilities p_{kj}, and the numbers n_k of vertices, $1 \leq k \leq K$, and m_j of objects, $1 \leq j \leq J$, of the different types.

The structure of the paper is as follows. The link between intersection graphs and branching processes is described in Section 2. The necessary distributional properties of the branching process are established in Section 3, and the extent to which it differs from the intersection graph is controlled in Section 4. The main theorem is then stated and proved in Section 6, and an application to exponential random graph models is given in Section 7. A key element in the proof is a Poisson approximation to coincidence probabilities in a generalization of the hypergeometric sampling scheme; this is undertaken in Section 5.

Although our motivation for studying this problem comes from the bipartite setting, one could equally well conduct a similar analysis for a graph without bipartite structure, recovering the general Erdős–Rényi mixture model; a corresponding approximation is given without detailed proof in (6.31). However, the analysis for a 'general' graph would not easily imply our results as a special case, with the vertices split into two groups and with a bipartite matrix of edge probabilities P, because the 2-periodic structure would result in there being more than one eigenvalue of the mean matrix having largest modulus, and methods such as those of this paper would still be needed, to deal with the extra complexity that results.

2 Intersection graphs and branching processes

A random multitype intersection graph on the vertex set $V = V_1 \uplus \ldots \uplus V_K$ is defined using a second set of 'objects' $U = U_1 \uplus \ldots \uplus U_J$: each vertex $v \in V_k$ independently chooses a subset $A_v \in U$ with distribution depending on k alone, and $v \sim v'$ if and only if $A_v \cap A_{v'} \neq \emptyset$. Here, we restrict ourselves to graphs derived from an underlying Erdős–Rényi bipartite mixture model, in which only edges e_{uv} between $u \in U$ and $v \in V$ are possible, and these are present or absent independently, with probability p_{kj} if $u \in U_j$ and $v \in V_k$. Thus, in the random intersection graph itself, $v \sim v'$ if and only if, for some $u \in U$, both e_{uv} and $e_{uv'}$ are present.

Such a random graph can be constructed from a bipartite multitype branching process $(Z(0), Z(1), Z(2), \ldots) = (X(0), Y(1), X(1), \ldots)$, with $X(i) \in \mathbf{N}^K$ and $Y(i) \in \mathbf{N}^J$ for each i, together with sets of randomly assigned indices. Start with numbers $X(0) = (X_1(0), \ldots, X_K(0))$ of individuals of the different types $\{(k,1),\, 1 \le k \le K\}$. The s-th individual (in some ordering) of type $(k,1)$ in $Z(2r) = X(r)$ has offspring vector $Y_{ks;r} = (Y^{(1)}_{ks;r}, \ldots, Y^{(J)}_{ks;r})$, realized from the product distribution $\otimes_{j=1}^J \mathrm{Bi}\,(m_j, p_{kj})$, where m_j is the cardinality of the set U_j, and the random vectors $(Y_{ks;r},\, 1 \le k \le K, s \ge 1, r \ge 0)$ are independent. We then set

$$Z(2r+1) \;=\; Y(r+1) \;=\; \sum_{k=1}^K \sum_{s=1}^{X_k(r)} Y_{ks;r}. \tag{2.1}$$

Similarly, the t-th individual of type $(j,2)$ in $Z(2r-1) = Y(r)$ has offspring vector $X_{jt;r} = (X^{(1)}_{jt;r}, \ldots, X^{(K)}_{jt;r})$, realized from the product distribution $\otimes_{k=1}^K \mathrm{Bi}\,(n_k, p_{kj})$, where n_k is the cardinality of the set V_k, and the random vectors $(X_{jt;r},\, 1 \le j \le J, s \ge 1, r \ge 1)$ are independent of each other and of the $Y_{ks;r}$. We then set

$$Z(2r) \;=\; X(r) \;=\; \sum_{j=1}^J \sum_{t=1}^{Y_j(r)} X_{jt;r}.$$

We also define

$$m \;:=\; \sum_{j=1}^J m_j; \qquad n \;:=\; \sum_{k=1}^K n_k. \tag{2.2}$$

Throughout this paper we assume that $m_j \geq 2$ for $1 \leq j \leq J$, and $n_k \geq 2$ for $1 \leq k \leq K$, and that $X_k(0) \leq n_k$ for $1 \leq k \leq K$.

To obtain the intersection graph, label each individual in the bipartite branching process with its line of descent. Thus

$$\{i; (k_0, s_0), (j_1, t_1), (k_1, s_1), \ldots, (j_i, t_i), (k_i, s_i)\}$$

labels the s_i-th individual of type $(k_i, 1)$ in generation $2i$, which was descended from the t_i-th individual of type $(j_i, 2)$ in generation $2i - 1$, and so on. These labels are then augmented with indices from the index set appropriate to the type of individual, as follows. The $Y^{(j)}_{k_i s_i; i}$ type $(j, 2)$ offspring of the typical vertex above are each assigned at random a unique index from a uniformly and independently chosen subset $\mathcal{L}^{(j)}_{k_i s_i; i} \subset \{1, 2, \ldots, m_j\}$ of size $Y^{(j)}_{k_i s_i; i}$; a similar construction is used for the offspring of objects.

A further class identifier, 0 or 1, is then attached to each individual: 1 if the individual is in generation zero, and thereafter, taking the individuals of the bipartite process in order of generation, but in any order within each generation, assign class 0 if its parent was in class 0, or if its index had previously been assigned to another individual of the same type and of class 1, and 1 otherwise. The class 0 individuals we refer to as ghosts. Edges are also created between the indices of a parent and its child if both are of class 1, or if the parent is of class 1 and the child of class $0'$, where class $0'$ indicates a class 0 individual whose index was first assigned (therefore to a class 1 individual) in its own generation. Then the individuals that belong to class 1 correspond, via their indices, to the vertices and objects used in constructing the intersection graph, and two vertices have an edge between them if there are corresponding class 1 or class $0'$ individuals in the bipartite branching process that are at distance 2 from one another. In this way, the union of the components of the random intersection graph that contain the initial vertices is sequentially constructed according to distance from the initial vertices, the class 1 vertices in generation $2i$ of the bipartite branching process corresponding to the vertices that are at distance i in the random intersection graph from the initially chosen set of vertices. If these components do not exhaust all vertices, the process can be continued from any unused vertex, until all have been used.

Two vertices, A of type k_1 and B of type k_2, are at distance at least $d + 1$ from one another in the random intersection graph if the d-neighbourhood of one of them in the bipartite graph does not intersect the d-neighbourhood of the other. Constructing the random intersection graph from a bipartite branching process starting with one individual A of type $(k_1, 1)$ and one, B, of type $(k_2, 1)$, this is the case exactly on the event that the set of all class 1 or class $0'$ descendants of A up to time d — both of types $(k, 1)$ and of types $(j, 2)$ — is disjoint from that of B. From the construction, these sets can only overlap if there is at least one class 1 descendant of either A or B having the same index as a class $0'$ descendant of the other, and then necessarily in the same generation of the bipartite process. Our main theorem consists of showing that the probability of this event can be well approximated by the probability of the corresponding event when *all* descendants are considered, and that this probability in turn can be well approximated using the theory of branching processes.

The origin of the approximation lies in the following well known facts ([10], II Theorem 9.2): that, on the event of non-extinction, a square integrable super-critical multitype branching process, whose mean matrix is irreducible and aperiodic, has an asymptotically stable type distribution; and that the total number of individuals alive in each generation grows like a random multiple of a geometrically growing sequence. For the X branching process, this means that the number of individuals of type k in generation i is approximately given by $\tau^i W \mu_k$, where τ is the largest eigenvalue of the mean matrix, μ^T is the associated positive left eigenvector, and W is a non-negative random variable, positive on the event of non-extinction, and the same for all i and k. Hence the numbers of descendants $X^A(i)$ of A at the i-th generation of the X branching process are approximately given by $\tau^i W^A \mu_k$, and those of B by $\tau^i W^B \mu_k$, where W^A and W^B are independent. When constructing the random intersection graph from the branching process, indices are assigned to the vertices independently at random, with replacement. Links between the A and B neighbourhoods occur whenever, for some $i \geq 1$ and some $1 \leq k \leq K$, one or more of the $X_k^B(i)$ are assigned the same index as one of the $X^A(i)$; other coincidences give rise to 'ghosts', and play no part in the intersection graph. The mean number of such events up to and including generation i is thus approximately

$$\sum_{s=1}^{i} \tau^{2s} W^A W^B \sum_{k=1}^{K} \frac{\mu_k^2}{n_k} \;\approx\; \kappa^X n^{-1} \tau^{2i} W^A W^B,$$

where

$$\kappa^X \;:=\; \left(\frac{\tau^2}{\tau^2 - 1}\right) \sum_{k=1}^{K} \frac{\mu_k^2}{q_k^X},$$

and $q_k^X := n_k/n$. A similar formula hold for links occurring because of coincidence of indices at the object level; here, the expected number of links up to and including generation $i-1$ is approximately $\kappa^Y n^{-1} \tau^{2(i-1)} W^A W^B$, where, because of (3.3), $\kappa^Y = \tau\kappa^X$. Adding the two means gives an overall mean number of links approximately equal to $\kappa n^{-1} \tau^{2i} W^A W^B$, where

$$\kappa \;:=\; \kappa^X (1 + \tau^{-1}) \;=\; \frac{\tau}{\tau - 1} \sum_{k=1}^{K} \frac{\mu_k^2}{n_k}. \qquad (2.3)$$

Then, using Poisson approximation, it follows that the probability of there being no shared vertices in the i-neighbourhoods of A and B is approximately

$$\mathbf{E}_{k_1,k_2} \left\{ e^{-\kappa n^{-1} \tau^{2i} W^A W^B} \right\},$$

this being the probability that the distance between A and B in the intersection graph exceeds $2i$. This line of reasoning is made precise in the coming sections, and the detailed results are to be found in Theorem 6.2 and Corollary 6.4.

3 Counting the offspring

We now study the bipartite branching process Z in greater detail. Our aim in this section is to justify the simple approximation, outlined above, to the numbers $X_k(i)$ of type-$(k, 1)$

individuals in $Z(2i)$ (or, equivalently, of type-k vertices in the i-th generation of the vertex branching process) and $Y_j(i)$ of type-$(j, 2)$ individuals in $Z(2i-1)$, (or of objects in the i-th generation of the object branching process). Theorems 3.6 and 3.7 below show that, for large i, $X(i) \sim \tau^i W\mu$ and $Y(i) \sim \tau^{i-1}\zeta W\tilde{\mu}$, the notation being as defined below.

3.1 Assumptions and notation

Let

$$N_X \;=\; \mathrm{diag}\,(n_1, \ldots, n_K) \quad \text{and} \quad N_Y \;=\; \mathrm{diag}\,(m_1, \ldots, m_J)$$

be the diagonal matrices of the numbers of different types of vertices, and of different types of objects, respectively; for future convenience, recalling (2.2), we define

$$q_k^X \;:=\; \frac{n_k}{n}, \quad 1 \le k \le K; \qquad q_j^Y \;:=\; \frac{m_j}{m}, \quad 1 \le j \le J. \tag{3.1}$$

Let P denote the $K \times J$-matrix of edge probabilities, and put

$$M_X \;=\; PN_YP^TN_X.$$

Then the non-negative matrix M_X is the mean matrix of the X branching process.

Assumption. We assume that the non-negative matrix M_X is irreducible and aperiodic, and has largest eigenvalue $\tau > 1$.

We use ν and μ^T to denote respectively the right and left eigenvectors corresponding to τ, with $\mu_k > 0$, $1 \le k \le K$, standardized so that $\|\mu\|_1 = 1$ and that $\mu^T\nu = 1$. We assume throughout that $\tau > 1$. We then define $M_Y := P^TN_XPN_Y$ to be the mean matrix of the Y branching process, and

$$\tilde{\mu} := \zeta^{-1}N_YP^T\mu \tag{3.2}$$

to be the left eigenvector of M_Y with $\|\tilde{\mu}\|_1 = 1$ corresponding to the eigenvalue τ. Thus $\zeta := \mu^TPN_Y\mathbf{1}$, where $\mathbf{1}$ denotes a $J \times 1$-vector of 1's.

Note also that

$$\begin{aligned}\zeta^2 \sum_{j=1}^{J} \frac{\tilde{\mu}_j^2}{m_j} \;&=\; (N_YP^T\mu)^TN_Y^{-1}(N_YP^T\mu) \;=\; \mu^TPN_YP^T\mu \\ &=\; \mu^TM_XN_X^{-1}\mu \;=\; \tau\mu^TN_X^{-1}\mu \;=\; \tau\sum_{k=1}^{K}\frac{\mu_k^2}{n_k},\end{aligned}$$

so that

$$\zeta^2 n/m \;=\; \tau\left\{\sum_{k=1}^{K}\frac{\mu_k^2}{q_k^X}\right\} \Big/ \left\{\sum_{j=1}^{J}\frac{\tilde{\mu}_j^2}{q_j^Y}\right\} \;=:\; \mathfrak{z}^2, \tag{3.3}$$

say.

We next define c_0 to be the smallest value such that

$$\begin{aligned} \sup_{a:\,\|a\|_1=1} \sup_{i\geq 0} \tau^{-i} a^T M_X^i e^{(k)} &\leq c_0 \mu_k, \qquad 1 \leq k \leq K; \\ \sup_{a:\,\|a\|_1=1} \sup_{i\geq 0} \tau^{-i} a^T M_Y^i \tilde{e}^{(j)} &\leq c_0 \tilde{\mu}_j, \qquad 1 \leq j \leq J, \end{aligned} \tag{3.4}$$

where $e^{(k)}$ and $\tilde{e}^{(j)}$ denote the k and j unit vectors in $\mathbf{R}^K$ and $\mathbf{R}^J$ respectively. Further, with λ_2 the eigenvalue of M_X with second largest modulus, we define c_1 such that

$$\sup_{b:\,\|b\|_\infty=1} \sup_{i\geq 0} |\lambda_2|^{-i} \|(M_X - \tau\nu\mu^T)^i b\|_\infty \leq c_1, \tag{3.5}$$

where we take $(M_X - \tau\nu\mu^T)^0 := I - \nu\mu^T$. Note that it follows from the Perron-Frobenius Theorem that c_0 and c_1 are both finite; see [11], Theorem 8.5.1, and [15], Chapter 1, Theorem 6.1. We also, for later use, write

$$\gamma := \max\{\tau, |\lambda_2|^2\} < \tau^2; \qquad \theta := \max_{s\geq 0}(s+1)(\sqrt{\gamma}/\tau)^s, \tag{3.6}$$

and introduce the notation $\widetilde{\mathcal{F}}_r$ to denote the σ-algebra $\sigma\{Z(t),\, 0 \leq t \leq r\}$, and $\mathcal{F}_i^X := \sigma\{X(l),\, 0 \leq l \leq i\}$, $\mathcal{F}_i^Y := \sigma\{Y(l),\, 1 \leq l \leq i\}$.

3.2 Asymptotics

The main results of the paper require no particular asymptotic setting. However, asymptotics are useful for putting the results in the context of a natural limiting framework. One such choice is the following. Start by choosing the n_k and m_j so that the proportions $q_k^X(m,n)$ and $q_j^Y(m,n)$ converge to non-zero limits. Then one can arrange for $P = P^{(m,n)}$ to vary as m and n tend to infinity, in such a way that the matrix $M_X^{(m,n)}$ converges to a fixed irreducible and aperiodic matrix M_X, entailing the convergence of quantities such as $\tau^{(m,n)}$, $\mu^{(m,n)}$ and $\nu^{(m,n)}$ to limits μ, ν and τ. With this in mind, define $Q_X^{(m,n)} := n^{-1}N_X$ and $Q_Y^{(m,n)} := m^{-1}N_Y$, and take $P^{(m,n)} := (mn)^{-1/2}\Pi$, for a fixed matrix Π. This then gives $M_X^{(m,n)} = \Pi Q_Y^{(m,n)} \Pi^T Q_X^{(m,n)}$, so that, if $Q_X^{(m,n)} \to Q_X$ and $Q_Y^{(m,n)} \to Q_Y$, with Q_X and Q_Y both having positive diagonals, then $M_X^{(m,n)} \to M_X := \Pi Q_Y \Pi^T Q_X$. If also, in keeping with the general assumptions of the paper, we have $\tau > 1$, then we describe this behaviour as 'standard asymptotics'.

Other asymptotic settings could equally well be considered. For instance, there would be no great difference in the qualitative behaviour if $\tau^{(m,n)}$ were allowed to tend to infinity with n like a power of $\log n$.

3.3 Expectations

We begin our analysis by examining the growth of the mean numbers of individuals of different types. Using $\mathbf{E}_0$ to denote $\mathbf{E}\{\cdot \,|\, \widetilde{\mathcal{F}}_0\}$, we immediately have

$$\mathbf{E}\{X^T(i) \,|\, \widetilde{\mathcal{F}}_{2i-1}\} = Y^T(i) P^T N_X; \qquad \mathbf{E}_0 X^T(i) = X^T(0) M_X^i, \tag{3.7}$$

and

$$\mathbf{E}\{Y^T(i)\,|\,\widetilde{\mathcal{F}}_{2i-2}\} \;=\; X^T(i-1)PN_Y; \qquad \mathbf{E}_0 Y^T(i) \;=\; X^T(0)M_X^{i-1}PN_Y. \tag{3.8}$$

From these, and using (3.4), we have, for instance,

$$\begin{aligned}
\mathbf{E}\{X_k(i)\,|\,\mathcal{F}_s^Y\} \;&=\; \sum_{j=1}^{J} Y^T(s)M_Y^{i-s}\tilde{e}^{(j)}\,(\tilde{e}^{(j)})^T P^T N_X e^{(k)} \\
&\le\; c_0\|Y(s)\|_1\tau^{i-s}\tilde{\mu}^T P^T N_X e^{(k)} \;=\; c_0\|Y(s)\|_1\tau^{i-s}\zeta^{-1}\mu^T M_X e^{(k)} \\
&=\; c_0\|Y(s)\|_1\tau^{i-s}(\tau/\zeta)\mu_k,
\end{aligned}$$

so that, for any $1 \le k \le K$ and $1 \le j \le J$, and for $i \ge s \ge 0$,

$$\mathbf{E}\{X_k(i)\,|\,\mathcal{F}_s^X\} \;\le\; c_0\|X(s)\|_1\tau^{i-s}\mu_k; \;\; \mathbf{E}\{X_k(i)\,|\,\mathcal{F}_s^Y\} \;\le\; c_0\|Y(s)\|_1\zeta^{-1}\tau^{i-s+1}\mu_k; \tag{3.9}$$

$$\mathbf{E}\{Y_j(i)\,|\,\mathcal{F}_s^Y\} \;\le\; c_0\|Y(s)\|_1\tau^{i-s}\tilde{\mu}_j; \;\; \mathbf{E}\{Y_j(i)\,|\,\mathcal{F}_{s-1}^X\} \;\le\; c_0\|X(s-1)\|_1\zeta\tau^{i-s}\tilde{\mu}_j; \tag{3.10}$$

$$\mathbf{E}_0 X_k(i) \;\le\; c_0\|X(0)\|_1\tau^i\mu_k; \;\; \mathbf{E}_0 Y_j(i) \;\le\; c_0\|X(0)\|_1\zeta\tau^{i-1}\tilde{\mu}_j. \tag{3.11}$$

3.4 X-Covariances

Controlling the covariances of the components of $X(i)$ and $Y(i)$ requires more work. To start with, we observe that

$$\mathbf{E}(X^T(i+1)\nu\,|\,\mathcal{F}_i^X) \;=\; X^T(i)M\nu \;=\; \tau X^T(i)\nu,$$

so that $W_i := \tau^{-i}X^T(i)\nu$, $i \ge 0$, is a non-negative martingale with respect to the filtration $\{\mathcal{F}_i^X, i = 0, 1, \ldots\}$ which converges almost surely to a limit W, and

$$\mathbf{E}_0(X^T(i)\nu) \;=\; \tau^i X^T(0)\nu. \tag{3.12}$$

The variability in the branching process is essentially determined by that of W, which is itself largely determined during the early stages of development. Indeed, writing $X^T(i) = X^T(i)\{I - \nu\mu^T\} + X^T(i)\nu\mu^T$, we show that the variance of $X^T(i)b$ is dominated by $(\mu^T b)^2\mathrm{Var}\,(X^T(i)\nu)$, unless $\mu^T b = 0$.

Lemma 3.1 *The variance of the martingale W_i is bounded as follows:*

$$\begin{aligned}
&(i) \qquad \mathrm{Var}_0\,(W_i - W) \;\le\; c_2\{\tau/(\tau-1)\}\|X(0)\|_1\tau^{-i}; \\
&(ii) \qquad \mathrm{Var}_0 W_i \;\le\; c_2\{\tau/(\tau-1)\}\|X(0)\|_1,
\end{aligned}$$

where $c_2 := c_0\tau^{-2}\|\nu\|_\infty^2 S(\Sigma)$ and $S(\Sigma) := \max_{1\le k\le K}\sum_{l,m=1}^K |\Sigma_{lm}^{[k]}|$; $\Sigma^{[k]}$ is defined below.

Proof: We begin by writing

$$X^T(i+1) \;=\; \sum_{k=1}^{K}\sum_{r=1}^{X_k(i)} \widetilde{X}_{kr;i}^T,$$

where $\widetilde{X}_{kr;i}$ denotes the K-vector of descendants in X-generation $i+1$ of individual r of type k in X-generation i. The random vectors $(\widetilde{X}_{kr;i},\, 1 \le k \le K, r \ge 1, i \ge 0)$ are independent, and, for each k, the $\widetilde{X}_{kr;i}$ are identically distributed, with means the transpose $M_{X,[k]}$ of the k-th row of M_X, and with a covariance matrix that we denote by $\Sigma^{[k]}$. Then

$$X^T(i+1) \;=\; \sum_{k=1}^{K}\sum_{r=1}^{X_k(i)} M_{X,[k]}^T + \sum_{k=1}^{K}\sum_{r=1}^{X_k(i)} (\widetilde{X}_{kr;i}^T - M_{X,[k]}^T),$$

and so

$$X^T(i+1) - X^T(i)M_X \;=\; \sum_{k=1}^{K}\sum_{r=1}^{X_k(i)} (\widetilde{X}_{kr;i}^T - M_{X,[k]}^T). \tag{3.13}$$

Considering the right hand side, we have

$$\begin{aligned}
&\mathbf{E}\left\{ \left(\sum_{k=1}^{K}\sum_{r=1}^{X_k(i)} (\widetilde{X}_{kr;i} - M_{X,[k]}) \right) \left(\sum_{k'=1}^{K}\sum_{r'=1}^{X_{k'}(i)} (\widetilde{X}_{k'r';i} - M_{X,[k']}) \right)^T \,\middle|\, \mathcal{F}_i \right\} \\
&= \sum_{k=1}^{K}\sum_{r=1}^{X_k(i)} \mathbf{E}\{ (\widetilde{X}_{kr;i} - M_{X,[k]})(\widetilde{X}_{kr;i} - M_{X,[k]})^T \,|\, \mathcal{F}_i^X \} \\
&= \sum_{k=1}^{K} X_k(i)\Sigma^{[k]}, \qquad (3.14)
\end{aligned}$$

by the independence of the vectors and their having mean zero. Hence, using (3.14), it follows that

$$\mathbf{E}\left\{ \left(\tau^{-(i+1)}X^T(i+1)\nu - \tau^{-i}X^T(i)\nu \right)^2 \,\middle|\, \mathcal{F}_i \right\} \;=\; \tau^{-2(i+1)} \sum_{k=1}^{K} X_k(i)\nu^T\Sigma^{[k]}\nu, \tag{3.15}$$

and thus, from (3.11),

$$\begin{aligned}
\mathbf{E}_0\left\{ \left(\tau^{-(i+1)}X^T(i+1)\nu - \tau^{-i}X^T(i)\nu \right)^2 \right\} \;&\le\; \sum_{k=1}^{K} c_0\|X(0)\|_1\tau^{-i-2}\mu_k\nu^T\Sigma^{[k]}\nu \\
&\le\; c_0\|X(0)\|_1\tau^{-i-2}\|\nu\|_\infty^2 S(\Sigma) \;=\; c_2\|X(0)\|_1\tau^{-i}, \qquad (3.16)
\end{aligned}$$

where $S(\Sigma) := \max_{1\le k\le K}\sum_{l,m=1}^{K}|\Sigma_{lm}^{[k]}|$ and $c_2 := c_0\tau^{-2}\|\nu\|_\infty^2 S(\Sigma)$. Writing the martingale $\tau^{-i}X^T(i)\nu$ as a sum of its one-step differences, the lemma now follows easily. □

The next lemma controls the variances of those components of $X(i)$ that are orthogonal to ν; note that

$$X^T(i) \;=\; (X^T(i)\nu)\mu^T + X^T(i)(I - \nu\mu^T).$$

Lemma 3.2 *With c_2 as in Lemma 3.1 and $\gamma := \max\{\tau, |\lambda_2|^2\} < \tau^2$, and with $c_3 := c_1^2 S(\Sigma)$, we have*

$$(i) \qquad \mathrm{Var}_0\{X^T(i)(I-\nu\mu^T)b\} \ \le\ c_3\|X(0)\|_1\, i\gamma^i\, \|b\|_\infty^2;$$
$$(ii) \qquad \mathrm{Var}_0\{X^T(i)b\} \ \le\ 2\|X(0)\|_1\, \{c_2\{\tau/(\tau-1)\}\, \tau^{2i}\, (\mu^T b)^2 + c_3\, i\gamma^i\, \|b\|_\infty^2\}$$

for any $b \in \mathbf{R}^K$. In particular, with $b = e_k$ the k^{th} unit vector it follows that

$$(iii) \qquad \mathrm{Var}_0\{X_k(i) - (X^T(i)\nu)\mu_k\} \ \le\ c_3\|X(0)\|_1\, i\gamma^i, \qquad 1 \le k \le K.$$

PROOF: Recalling that $\mu^T\nu = 1$, it can be seen by induction that

$$\big(M_X - \tau(\nu\mu^T)\big)^i \ =\ M_X^i - \tau^i(\nu\mu^T). \tag{3.17}$$

Hence, with $M_X^0 = I$ and as $\nu\mu^T(M_X - \tau\nu\mu^T) = 0$,

$$\begin{aligned} & X^T(i)(I-\nu\mu^T) \\ &= \sum_{r=0}^{i-1} \big\{X^T(r+1)\left[M_X^{i-r-1} - (\nu\mu^T)\tau^{(i-r-1)}\right] - X^T(r)\left[M_X^{i-r} - (\nu\mu^T)\tau^{(i-r)}\right]\big\} \\ &\quad + X^T(0)\left[M_X^i - (\nu\mu^T)\tau^i\right] \\ &= \sum_{r=0}^{i-1} \big\{X^T(r+1)(I-\nu\mu^T) - X^T(r)(M_X - \tau(\nu\mu^T))\big\}\left[M_X - \tau(\nu\mu^T)\right]^{i-r-1} \\ &\quad + X^T(0)\left[M_X^i - (\nu\mu^T)\tau^i\right] \\ &= \sum_{r=0}^{i-1} U^T(r)A^{i-r-1} + X^T(0)\left[M_X^i - (\nu\mu^T)\tau^i\right], \end{aligned} \tag{3.18}$$

where

$$U^T(r) \ :=\ (X^T(r+1) - X^T(r)M_X)(I-\nu\mu^T) \quad \text{and} \quad A \ :=\ M_X - \tau(\nu\mu^T).$$

Note, in particular, that $\mathbf{E}_0\{U(r)U^T(s)\} = 0$ whenever $r \ne s$, in view of (3.13), and that

$$\mathbf{E}\{U(r)U^T(r)\,|\,\mathcal{F}_r^X\} \ =\ \sum_{k=1}^K X_k(r)(I-\mu\nu^T)\Sigma^{[k]}(I-\nu\mu^T). \tag{3.19}$$

Hence, since $(I-\nu\mu^T)(M_X - \tau\nu\mu^T) = M_X - \tau\nu\mu^T$, it follows from (3.5) and (3.11) that

$$\begin{aligned} \mathrm{Var}_0\{X^T(i)(I-\nu\mu^T)b\} \ &=\ b^T\sum_{r=0}^{i-1}(A^T)^{i-1-r}\mathbf{E}_0\{U(r)U^T(r)\}A^{i-1-r}b \\ &\le\ \sum_{r=0}^{i-1}\sum_{k=1}^K \mathbf{E}_0 X_k(r)\{c_1|\lambda_2|^{(i-1-r)}\|b\|_\infty\}^2 S(\Sigma) \\ &\le\ c_0\|X(0)\|_1\sum_{r=0}^{i-1}\tau^r\{c_1|\lambda_2|^{(i-1-r)}\|b\|_\infty\}^2 S(\Sigma), \end{aligned} \tag{3.20}$$

for any $b \in \mathbf{R}^K$, proving part (i), and part (iii) follows directly. Since also, from Lemma 3.1 (ii),

$$\mathrm{Var}_0\{X^T(i)\nu\mu^T b\} \;\le\; c_2\{\tau/(\tau-1)\}\|X(0)\|_1\tau^{2i}(\mu^T b)^2, \tag{3.21}$$

it follows from part (i) that

$$\begin{aligned}\mathrm{Var}_0\{X^T(i)b\} &\le 2\big\{\mathrm{Var}_0\{X^T(i)\nu\mu^T b\} + \mathrm{Var}_0\{X^T(i)(I-\nu\mu^T)b\}\big\} \\ &\le 2\{c_2\{\tau/(\tau-1)\}\|X(0)\|_1\,\tau^{2i}\,(\mu^T b)^2 + c_3\|X(0)\|_1\, i\gamma^i\,\|b\|_\infty^2\},\end{aligned} \tag{3.22}$$

establishing part (ii). □

Note that the growth of $\mathrm{Var}_0\{X^T(i)b\}$ with i is at rate $O(i\gamma^i\,\|b\|_\infty^2)$, slower than τ^{2i}, if $\mu^T b = 0$. Note also that, if $\tau \ne |\lambda_2|^2$, the factor i can be replaced by a constant $c(\tau, |\lambda_2|^2)$.

Corollary 3.3 *For all $1 \le k, l \le K$, there are constants c_5, c_6 such that*

$$\mathbf{E}_0\{X_k(i)X_l(i)\} \;\le\; c_5\|X(0)\|_1^2\mu_k\mu_l\tau^{2i} + c_6\|X(0)\|_1 i\tau^i\gamma^{i/2}.$$

In particular, for $c_5' := c_5 + c_6K^2\theta$, where θ is as in (3.6), we have

$$\mathbf{E}_0\|X(i)\|_1^2 \;\le\; c_5'\tau^{2i}\|X(0)\|_1^2.$$

Proof: It follows from (3.11) and Lemma 3.2 (ii) with the Cauchy-Schwarz inequality that

$$\begin{aligned}\mathbf{E}_0\{X_k(i)X_l(i)\} &\le |\mathbf{E}_0X_k(i)\mathbf{E}_0X_l(i)| + |\mathrm{Cov}_0(X_k(i),X_l(i))| \\ &\le \{c_0^2\|X(0)\|_1^2 + 2c_2\|X(0)\|_1\tau/(\tau-1)\}\mu_k\mu_l\tau^{2i} \\ &\quad + 4\|X(0)\|_1(i\gamma^i)^{1/2}\tau^i\sqrt{c_2c_3\{\tau/(\tau-1)\}} + 2c_3\|X(0)\|_1 i\gamma^i,\end{aligned} \tag{3.23}$$

and the corollary follows by taking $c_5 := c_0^2 + 2c_2\tau/(\tau-1)$ and $c_6 := 4\sqrt{c_2c_3\{\tau/(\tau-1)\}} + 2c_3$. □

3.5 Y-Covariances

Very similar arguments can also be carried through for the vectors $Y(i)$, $i \ge 1$. We first show that $Y_j(i)$ is close enough to $X^T(i-1)PN_Y\tilde{e}^{(j)}$, for any $1 \le j \le J$. Let $Z_j(i) := Y_j(i) - X^T(i-1)PN_Y\tilde{e}^{(j)}$.

Lemma 3.4 *There is a constant c_{10} such that, for $1 \le j \le J$,*

$$\mathrm{Var}_0\{Z_j(i)\} \;\le\; c_{10}\|X(0)\|_1\tau^i\zeta\tilde{\mu}_j. \tag{3.24}$$

PROOF: First, the quantity $Z^T(i) := Y^T(i) - X^T(i-1)PN_Y$ is represented in fashion analogous to (3.13). For independent random vectors $Y_{ks,r}$ as in (2.1),

$$Z^T(i) = Y^T(i) - X^T(i-1)PN_Y = \sum_{k=1}^{K} \sum_{s=1}^{X_k(i-1)} (Y_{ks,i-1} - \mathbf{E}Y_{ks,i-1}),$$

from which it follows that

$$\mathbf{E}\{Z(i)Z^T(i) \,|\, \mathcal{F}_{i-1}^X\} \;=\; \sum_{k=1}^{K} X_k(i-1)\widetilde{\Sigma}^{[k]},$$

where $\widetilde{\Sigma}^{[k]} := \mathrm{diag}\,\{m_j p_{kj}(1-p_{kj}),\, 1 \le j \le J\}$, and hence that

$$\mathbf{E}_0\{Z(i)Z^T(i)\} \;=\; \sum_{k=1}^{K} \mathbf{E}_0 X_k(i-1)\widetilde{\Sigma}^{[k]}.$$

This with (3.11) and (3.2) in turn yields

$$\begin{aligned} \mathrm{Var}_0\{Z_j(i)\} \;&=\; \sum_{k=1}^{K} \mathbf{E}_0 X_k(i-1)\tilde{e}^{(j)T}\widetilde{\Sigma}^{[k]}\tilde{e}^{(j)} \\ &\le\; c_0\|X(0)\|_1 \tau^{i-1} \sum_{k=1}^{K} \mu_k m_j p_{kj}(1-p_{kj}) \;\le\; c_{10}\|X(0)\|_1 \zeta \tau^i \tilde{\mu}_j, \end{aligned} \tag{3.25}$$

with $c_{10} := c_0/\tau$. □

We now, for future use, define the quantity

$$\zeta_* \;:=\; J \max_{1\le j\le J} \|PN_Y \tilde{e}^{(j)}\|_\infty, \tag{3.26}$$

noting also that

$$\big(J^{-1}\min_k \mu_k\big)\zeta_* \;\le\; \zeta \;=\; \sum_{j=1}^{J} \mu^T PN_Y \tilde{e}^{(j)} \;\le\; \|\mu\|_1 \zeta_* \;=\; \zeta_*. \tag{3.27}$$

Hence, in view of (3.3), we write

$$\zeta_* \;:=\; Z_* \sqrt{\frac{m}{n}}, \tag{3.28}$$

noting that $\mathfrak{Z}$ and Z_* can be thought of as having comparable magnitude. We also introduce the notation

$$u_{mn} \;:=\; \left(\frac{m}{n}\right)^{1/4} \left\{1 + \left(\frac{m}{n}\right)^{1/4}\right\}. \tag{3.29}$$

Using Lemma 3.4, we can now establish an analogue of Corollary 3.3 for the elements of $Y(i)$.

Corollary 3.5 *For all $1 \le j,l \le J$, there is a constant c_6' such that*

$$\mathbf{E}_0\{Y_j(i)Y_l(i)\} \;\le\; c_5\zeta^2\|X(0)\|_1^2\tilde\mu_j\tilde\mu_l\tau^{2(i-1)} + c_6' u_{mn}^2\|X(0)\|_1 i\tau^{i-1}\gamma^{(i-1)/2},$$

where c_5 is as in Corollary 3.3. In particular, for $c_5'' := \mathfrak{Z}^2 c_5 + c_6' J^2\theta$, we have

$$\mathbf{E}_0\|Y(i)\|_1^2 \;\le\; c_5'' u_{mn}^2\tau^{2(i-1)}\|X(0)\|_1^2.$$

Proof: We first note that

$$\begin{aligned}\mathbf{E}_0\{Y_j(i)Y_l(i)\} &\le |\mathbf{E}_0 Y_j(i)\mathbf{E}_0 Y_l(i)| + |\mathrm{Cov}_0(Y_j(i),Y_l(i))| \\ &\le |\mathbf{E}_0 Y_j(i)\mathbf{E}_0 Y_l(i)| + \sqrt{\mathrm{Var}_0\{Y_j(i)\}\,\mathrm{Var}_0\{Y_l(i)\}}.\end{aligned}$$

Now, writing $Y_j(i) = X^T(i-1)PN_Y\tilde e^{(j)} + Z_j(i)$, it follows from Lemmas 3.2 (ii) and 3.4 and from (3.26) that

$$\begin{aligned}&\sqrt{\mathrm{Var}_0\{Y_j(i)\}} \\ &\quad\le \sqrt{2\|X(0)\|_1}\left\{\sqrt{c_2\tau/(\tau-1)}\tau^{i-1}\zeta\tilde\mu_j + \sqrt{c_3(i-1)\gamma^{i-1}}J^{-1}\zeta_* + \sqrt{c_{10}\zeta\tau^i/2}\right\} \\ &\quad\le \sqrt{2\|X(0)\|_1}\left\{\sqrt{c_2\tau/(\tau-1)}\tau^{i-1}\zeta\tilde\mu_j + u_{mn}\sqrt{c_4 i\gamma^{i-1}}\right\},\end{aligned}$$

where $\sqrt{c_4} = (J^{-1}\sqrt{c_3}Z_* \vee \sqrt{c_{10}\mathfrak{Z}\tau/2})$. A similar bound holds also for $\sqrt{\mathrm{Var}_0\{Y_l(i)\}}$. The corollary now follows from (3.11), and by taking $c_6' = 4\mathfrak{Z}\sqrt{c_2c_4\tau/(\tau-1)} + 2c_4$. □

3.6 X- and Y-approximation

Using the preparation above, we are now able to approximate $X_k(i)$ and $Y_j(i)$, $i \ge 1$, in terms of the limiting random variable W, making precise the description at the end of Section 2, and bounding the error in the approximation. We begin by considering the X-components.

Theorem 3.6 *There is a constant c_9 such that, for $1 \le k \le K$,*

$$\mathbf{E}_0\Big|X_k(i) - \tau^i W\mu_k\Big| \;\le\; c_9\|X(0)\|_1((i+1)\gamma^i)^{1/2}, \qquad i \ge 0. \tag{3.30}$$

Proof: Here, we note from Lemma 3.2 (iii), (3.5), (3.7) and (3.17) that

$$\begin{aligned}&\mathbf{E}_0\Big|X_k(i) - (X^T(i)\nu)\mu_k\Big| \\ &\quad\le \big(\mathrm{Var}_0\{X_k(i) - (X^T(i)\nu)\mu_k\}\big)^{1/2} + |\mathbf{E}_0\{X_k(i) - (X^T(i)\nu)\mu_k\}| \\ &\quad\le (c_3\|X(0)\|_1 i\gamma^i)^{1/2} + \|X(0)\|_1\|(M_X^i - \tau^i\nu\mu^T)e^{(k)}\|_\infty \\ &\quad\le c_7\|X(0)\|_1((i+1)\gamma^i)^{1/2},\end{aligned}$$

with $c_7 := \sqrt{c_3} + c_1$, whereas, from Lemma 3.1 (i), and the Cauchy-Schwarz inequality,

$$\mathbf{E}_0 \Big| (X^T(i)\nu - \tau^i W)\mu_k \Big| \;\le\; \mu_k \tau^i \{ c_2 \{\tau/(\tau-1)\} \|X(0)\|_1 \tau^{-i} \}^{1/2} \;\le\; c_8 \|X(0)\|_1 \tau^{i/2}, \quad (3.31)$$

with $c_8 := \{c_2\tau/(\tau-1)\}^{1/2}$. Hence the theorem follows, with $c_9 \;:=\; c_7 + c_8$. □

With the help of Lemma 3.4, we can also approximate $Y(i)$ in terms of the limiting random variable W, complementing Theorem 3.6.

Theorem 3.7 *There is a constant c_{14} such that, for each $1 \le j \le J$,*

$$\mathbf{E}_0 \Big| Y_j(i) - \tau^{i-1} \zeta W \tilde{\mu}_j \Big| \;\le\; c_{14} u_{mn} \|X(0)\|_1 (i\gamma^i)^{1/2}, \qquad i \ge 1, \qquad (3.32)$$

where $\tilde{\mu}$ is given in (3.2).

Proof: It is immediate from Lemma 3.4 that

$$\mathbf{E}_0 \Big| Y_j(i) - X^T(i-1) P N_Y \tilde{e}^{(j)} \Big| \;\le\; \{ c_{10} \zeta \|X(0)\|_1 \|\tilde{\mu}\|_\infty \tau^i \}^{1/2} \;\le\; c_{11} \zeta^{1/2} \|X(0)\|_1 \tau^{(i-1)/2},$$

with $c_{11} := \sqrt{c_{10}\|\tilde{\mu}\|_\infty}$, and then, as in the proof of Theorem 3.6, using Lemma 3.1 (i), (3.17) and (3.5), we have

$$\begin{aligned} &\mathbf{E}_0 \Big| X^T(i-1) P N_Y \tilde{e}^{(j)} - \zeta X^T(i-1)\nu \tilde{\mu}_j \Big| \\ &\quad \le\; \Big\{ (c_3 \|X(0)\|_1 (i-1)\gamma^{i-1})^{1/2} + c_1 \|X(0)\|_1 \gamma^{(i-1)/2} \Big\} \|P N_Y \tilde{e}^{(j)}\|_\infty \\ &\quad \le\; c_{12} \zeta_* \|X(0)\|_1 (i\gamma^i)^{1/2}, \end{aligned}$$

with $c_{12} := J^{-1}\sqrt{1/\gamma}(\sqrt{c_3} + c_1)$. Then, once again invoking Lemma 3.1 (i), as for (3.31), we have

$$\mathbf{E}_0 \Big| \zeta X^T(i-1)\nu\tilde{\mu}_j - \tau^{i-1}\zeta W \tilde{\mu}_j \Big| \;\le\; c_8 \zeta \tilde{\mu}_j \|X(0)\|_1 \tau^{(i-1)/2} \;\le\; c_{13} \zeta \|X(0)\|_1 \tau^{i/2},$$

with $c_{13} := (c_8/\sqrt{\tau})\|\tilde{\mu}\|_\infty$. The theorem follows, since $\gamma \ge \tau$, by taking $c_{14} := (\{c_{11}\sqrt{\mathfrak{Z}}\} \vee \{c_{12} Z_* + c_{13}\mathfrak{Z}\})$. □

4 Ghosts

In the previous section, we justified simple approximations to the joint counts $X(i)$ and $Y(i)$ in the bipartite branching process. We now need to show that the same approximation can be used for the composition of the neighbourhoods in the intersection graph, albeit with a further error. This involves showing that the effect of the 'ghosts' is not too large. Let $G_k^X(i)$ and $G_j^Y(i)$ denote the total numbers of type $(k,1)$ and of type $(j,2)$ individuals of class 0 (ghosts), respectively, alive in generations $2i$ and $2i-1$ respectively of the bipartite process that starts with individuals A and B. Then it turns out to be enough to derive bounds for their expectations, as functions of i.

To state the result, define

$$\rho^X \;:=\; \max_{1\le k\le K} \mu_k/q_k^X; \qquad \rho^Y \;:=\; \max_{1\le j\le J} \tilde\mu_j/q_j^Y, \tag{4.1}$$

where q^X and q^Y are as in (3.1). Note that, for the Z-process under consideration, $\|X(0)\|_1 = 2$.

Theorem 4.1 *There are constants c_{15}^* and c_{16}^* such that*

$$\begin{aligned}
\mathbf{E}_0 G_k^X(i) &\le c_{15}^* \tau^{2i} e(m,n)^4, \quad 1\le k\le K;\\
\mathbf{E}_0 G_j^Y(i) &\le c_{16}^* \sqrt{\frac{m}{n}}\,\tau^{2(i-1)} e(m,n)^4, \quad 1\le j\le J,
\end{aligned}$$

where $e(m,n) := n^{-1/4} + m^{-1/4}$.

Proof: The ghosts can be counted by descent from original ghosts, whose parents were of class 1. We write $H_k^X(l)$ and $H_j^Y(l)$ to denote the numbers of original ghosts of the corresponding types born in the l-th generations of the X and Y processes. For $i > l \ge 0$, we then let $G_k^{XY}(l,i)$ and $G_k^{XX}(l,i)$ denote the total numbers of descendants of generation l original Y and X ghosts, respectively, alive at time $2i$ in the bipartite process, that are $(k,1)$ individuals; the quantities $G_j^{YY}(l,i)$ and $G_j^{YX}(l,i)$ are defined analogously. Thus

$$G_k^X(i) = H_k^X(i) + \sum_{l=0}^{i-1}\{G_k^{XY}(l+1,i) + G_k^{XX}(l,i)\}; \tag{4.2}$$

$$G_j^Y(i) = H_j^Y(i) + \sum_{l=1}^{i-1}\{G_j^{YY}(l,i) + G_j^{YX}(l,i)\} + G_j^{YX}(0,i). \tag{4.3}$$

Note that, for $i \ge l \ge 0$, from (3.9) and (3.3),

$$\mathbf{E}(G_k^{XY}(l,i)\,|\,H^Y(l)) \;\le\; c_0 \mathfrak{Z}^{-1}\tau\|H^Y(l)\|_1 \tau^{i-l}\mu_k\sqrt{\frac{n}{m}}; \tag{4.4}$$

and that, for $i > l \ge 0$,

$$\mathbf{E}(G_k^{XX}(l,i)\,|\,H^X(l)) \;\le\; c_0\|H^X(l)\|_1\tau^{i-l}\mu_k; \tag{4.5}$$

$$\mathbf{E}(G_j^{YY}(l,i)\,|\,H^Y(l)) \;\le\; c_0\|H^Y(l)\|_1\tau^{i-l}\tilde\mu_j; \tag{4.6}$$

$$\mathbf{E}(G_j^{YX}(l,i)\,|\,H^X(l)) \;\le\; c_0\mathfrak{Z}\|H^X(l)\|_1\tau^{i-l-1}\tilde\mu_j\sqrt{\frac{m}{n}}, \tag{4.7}$$

from (3.9) and (3.10).

Now an original ghost of type $(j,2)$ is created when an index from the set $\{1,2,\ldots,m_j\}$ is re-used. Hence

$$\mathbf{E}\{H_j^Y(l)\,|\,\mathcal{F}_l^Y\} \;\le\; Y_j(l)\left\{m_j^{-1}\sum_{s=1}^{l} Y_j(s)\right\}.$$

Furthermore, from (3.10), for $l > s \geq 0$,

$$\mathbf{E}\{Y_j(l) \,|\, \mathcal{F}_s^Y\} \;\leq\; c_0 \|Y(s)\|_1 \tau^{l-s} \tilde\mu_j.$$

Combining these bounds, it follows that

$$\mathbf{E}_0 H_j^Y(l) \;\leq\; c_0 m_j^{-1} \tilde\mu_j \sum_{s=1}^{l} \tau^{l-s} \mathbf{E}_0(Y_j(s)\|Y(s)\|_1), \tag{4.8}$$

and hence, using Corollary 3.5, that

$$\mathbf{E}_0\|H^Y(l)\|_1 \;\leq\; c_0 m^{-1} \rho^Y \sum_{s=1}^{l} \tau^{l-s} \mathbf{E}_0 \|Y(s)\|_1^2 \;\leq\; c_{17} m^{-1}\, \rho^Y u_{mn}^2 \tau^{2(l-1)}, \tag{4.9}$$

where $c_{17} := 4c_5'' c_0 \tau/(\tau-1)$. Similar calculations show that

$$\mathbf{E}_0 H_k^X(l) \;\leq\; c_0 n_k^{-1} \mu_k \sum_{s=0}^{l} \tau^{l-s} \mathbf{E}_0(X_k(s)\|X(s)\|_1), \tag{4.10}$$

and, with Corollary 3.3,

$$\mathbf{E}_0\|H^X(l)\|_1 \;\leq\; c_{18} n^{-1} \rho^X \tau^{2l}, \tag{4.11}$$

with $c_{18} := 4c_5' c_0 \tau/(\tau-1)$. For future reference, we note also that, in consequence, for any $s \geq 0$,

$$\begin{aligned} \sum_{l=1}^{s} \tau^{s-l} \mathbf{E}_0\|H^Y(l)\|_1 \;&\leq\; \frac{c_{17}\tau}{\tau-1} m^{-1} \rho^Y u_{mn}^2 \tau^{2(s-1)}; \\ \sum_{l=0}^{s} \tau^{s-l} \mathbf{E}_0\|H^X(l)\|_1 \;&\leq\; \frac{c_{18}\tau}{\tau-1} n^{-1} \rho^X \tau^{2s}. \end{aligned} \tag{4.12}$$

It now remains to take expectations in (4.2) and (4.3), using (4.4)–(4.7) and (4.12). For instance, for $\mathbf{E}_0 G_j^Y(i)$, we have

$$\sum_{l=1}^{i-1} \mathbf{E}_0 G_j^{YY}(l,i) \;\;\leq\;\; c_0 \tilde\mu_j \sum_{l=1}^{i-1} \tau^{i-l} \mathbf{E}_0 \|H^Y(l)\|_1 \;\leq\; c_{19} m^{-1} \rho^Y \tilde\mu_j u_{mn}^2 \tau^{2(i-1)},$$

with $c_{19} := c_0 c_{17}/(\tau-1)$; similar calculations yield

$$\sum_{l=0}^{i-1} \mathbf{E}_0 G_j^{YX}(l,i) \;\;\leq\;\; c_{20} n^{-1} \rho^X \tilde\mu_j \tau^{2(i-1)} \sqrt{\frac{m}{n}},$$

with $c_{20} := 3 c_0 c_{18} \tau/(\tau-1)$, and

$$\mathbf{E}_0 H_j^Y(i) \;\;\leq\;\; c_{17} m^{-1} \rho^Y u_{mn}^2 \tau^{2(i-1)},$$

from (4.9). Hence it follows that

$$\mathbf{E}_0 G_j^Y(i) \;\le\; \sqrt{\frac{m}{n}}\,\tau^{2(i-1)}\left(\frac{c_{16}\rho^X}{n} + \frac{\tilde{c}_{16}\rho^Y}{m}\left\{1+\left(\frac{m}{n}\right)^{1/4}\right\}^2\right),$$

with $c_{16} := c_{20}\|\tilde{\mu}\|_\infty$ and $\tilde{c}_{16} := c_{19}\|\tilde{\mu}\|_\infty + c_{17}$. A similar argument yields the bound

$$\mathbf{E}_0 G_k^X(i) \;\le\; \tau^{2i}\left(\frac{c_{15}\rho^X}{n} + \frac{\tilde{c}_{15}\rho^Y}{m}\left\{1+\left(\frac{m}{n}\right)^{1/4}\right\}^2\right),$$

with

$$c_{15} \;:=\; c_{18}\Big\{c_0(\tau-1)^{-1}\,\|\mu\|_\infty + 1\Big\}; \quad \tilde{c}_{15} \;:=\; c_0 c_{17} 3^{-1}(\tau-1)^{-1}\,\|\mu\|_\infty.$$

Recalling the definition (3.28) of Z_*, this completes the theorem, with

$$c_{15}^* \;:=\; 2(c_{15}\rho^X \vee \tilde{c}_{15}\rho^Y); \qquad c_{16}^* \;:=\; 2(c_{16}\rho^X \vee \tilde{c}_{16}\rho^Y).$$

□

5 The probability of a common label

Our next step is to establish a Poisson approximation for the probability of a coincidence in a random labelling problem. The underlying idea is to look at neighbourhoods of radii i_A and i_B of two initial vertices A and B; if they have no vertices in common, then the distance between A and B exceeds $i_A + i_B$. Whether two vertices in the neighbourhoods are the same can be thought of as a labelling problem, where the assignment of labels is almost uniform and at random. The result that we need is the following variant on the Poisson approximation to hypergeometric sampling.

Proposition 5.1 *For each l, $1 \le l \le L$, and for each r, $1 \le r \le R_l$, we independently draw a subset $\{J_{lr1},\ldots,J_{lrz_{lr}}\}$ of size z_{lr} from a fixed set W_l of size $w_l \ge 2$, with any subset equally likely to be drawn; define $z_l := \sum_{r=1}^{R_l} z_{lr}$. We then repeat the experiment independently, with subsets $\{J'_{lr1},\ldots,J'_{lrz'_{lr}}\}$ of sizes z'_{lr}, $1 \le l \le L$, $1 \le r \le R'_l$. For each $w \in W_l$ and for each l, r, s, r', s', define*

$$H(w,l;r,s;r',s') \;:=\; I[J_{lrs} = J'_{lr's'} = w].$$

Then, for fixed subsets $W_l^* \subset W_l$ of sizes w_l^*, $1 \le l \le L$, define

$$S \;:=\; S(z,z';w,w^*) \;:=\; \sum_{l=1}^{L}\sum_{w\notin W_l^*}\sum_{r=1}^{R_l}\sum_{r'=1}^{R'_l}\sum_{s=1}^{z_{lr}}\sum_{s'=1}^{z'_{lr'}} H(w,l;r,s;r',s')$$

to be the number of pairs of elements, one from each sample, that consist of two copies of the same element, not belonging to any of the W_l^*. Then

$$|\mathbf{P}[S(z,z';w,w^*)=0] - \exp(-\lambda(z,z',w,L))| \;\le\; B_1(z,z',w,L) + B_1^*(z,z',w,w^*,L),$$

where

$$\lambda(z,z',w,L) \;:=\; \sum_{l=1}^{L} \frac{z_l z_l'}{w_l}; \qquad B_1(z,z',w,L) \;:=\; 2\sum_{l=1}^{L} \frac{z_l + z_l'}{w_l}, \tag{5.1}$$

and

$$B_1^*(z,z',w,w^*,L) \;:=\; \sum_{l=1}^{L} \frac{z_l z_l' w_l^*}{w_l^2}. \tag{5.2}$$

PROOF: The indicators $H(w,l;r,s;r',s')$ are negatively related (for the definition see [4] Definition 2.1.1), as can be seen by constructing an explicit coupling very much as in [4], p.112. Setting

$$\widetilde{H}(l,r,r') \;:=\; \sum_{w\notin W_l^*} \sum_{s=1}^{z_{lr}} \sum_{s'=1}^{z'_{lr'}} H(w,l;r,s;r',s'),$$

the random variables $\widetilde{H}(l,r,r')$ are pairwise independent, and satisfy

$$\begin{aligned}
&\mathbf{E}\widetilde{H}(l,r,r') \;=\; \left(1 - \frac{w_l^*}{w_l}\right)\frac{z_{lr}z'_{lr'}}{w_l};\\
&0 \;\le\; \mathbf{E}\widetilde{H}(l,r,r') - \operatorname{Var}\widetilde{H}(l,r,r') \;\le\; \mathbf{E}\widetilde{H}(l,r,r')\left(\frac{z_{lr}+z'_{lr'}-1}{w_l-1}\right);\\
&S \;=\; \sum_{l=1}^{L}\sum_{r=1}^{R_l}\sum_{r'=1}^{R_l'} \widetilde{H}(l,r,r').
\end{aligned}$$

Hence, since $w_l \ge 2$ for all l, it follows that

$$0 \;\le\; \mathbf{E}S - \operatorname{Var}S \;\le\; 2\max_{l,r,r'}\left(\frac{z_{lr}+z'_{lr'}}{w_l}\right)\mathbf{E}S,$$

so that

$$1 - \frac{\operatorname{Var}S}{\mathbf{E}S} \;\le\; 2\sum_{l=1}^{L}\sum_{r=1}^{R_l}\frac{z_{lr}}{w_l} + 2\sum_{l=1}^{L}\sum_{r'=1}^{R_l'}\frac{z'_{lr'}}{w_l} \;\le\; B_1(z,z',w,L).$$

From this, and since

$$\sum_{l=1}^{L}\sum_{r=1}^{R_l}\sum_{r'=1}^{R_l'} \mathbf{E}\widetilde{H}(l,r,r') \;=\; \lambda(z,z',w,L) - B_1^*(z,z',w,w^*,L),$$

it follows using [4] Theorem 2.C.2 that

$$\left|\mathbf{P}[S(z,z';w,w^*)=0] - \exp\{-[\lambda(z,z',w,L) - B_1^*(z,z',w,w^*,L)]\}\right| \;\le\; B_1(z,z',w,L). \tag{5.3}$$

The final estimate follows because

$$|e^{-\lambda} - e^{-\lambda'}| \le \min\{1, |\lambda - \lambda'|\} \tag{5.4}$$

when $\lambda, \lambda' \ge 0$. □

Note that, if, for some $\tau > 1$, none of z_l and z'_l exceeds $C_1\tau^{i/2}$ and all of the w_l exceed $C_2\tau^i$, then

$$B_1(z, z', w, L) \;\le\; \frac{4LC_1}{C_2\,\tau^{i/2}},$$

geometrically small with i.

Now suppose that we do not know the true values z_l and z'_l, but only approximations $\tilde{z}_l$ and $\tilde{z}'_l$ to them. Then we can instead use these to approximate $\mathbf{P}[S(z, z', w, w^*) = 0]$, with some possible extra error.

Proposition 5.2 *Suppose that*

$$|\tilde{z}_l - z_l| = \varepsilon_l \quad \textit{and} \quad |\tilde{z}'_l - z'_l| = \varepsilon'_l, \qquad 1 \le l \le L.$$

Then

$$\begin{aligned}&|\mathbf{P}[S(z, z', w, w^*) = 0] - \exp(-\lambda(\tilde{z}, \tilde{z}', w, L))|\\ &\quad \le\; B_1(z, z', w, L) + B_1^*(z, z', w, w^*, L) + B_2(z, z', \varepsilon, \varepsilon', w, L),\end{aligned}$$

where

$$B_2(z, z', \varepsilon, \varepsilon', w, L) \;=\; \bar{\lambda}(z, \varepsilon', w, L) + \bar{\lambda}(\varepsilon, z', w, L) + \bar{\lambda}(\varepsilon, \varepsilon', w, L),$$

where $\bar{\lambda} := \min(\lambda, 1)$.

PROOF: Immediate from (5.4). □

In Section 6, we take for $\tilde{z}$ and $\tilde{z}'$ convenient approximations to numbers of individuals alive in generations $r \ge 1$ in the bipartite branching process that starts from the two individuals A of type $(k_1, 1)$ and B of type $(k_2, 1)$. For $r = 2l$ even, we take $\tilde{z}$ to approximate the $X^A(l)$ descendants of A, and $\tilde{z}'$ to approximate the $X^B(l)$ descendants of B, and $w_l = n_l$, $1 \le l \le K$. For $r = 2l - 1$ odd, we take $\tilde{z}$ to approximate the $Y^A(l)$ descendants of A, and $\tilde{z}'$ to approximate the $Y^B(l)$ descendants of B, now with $w_l = m_l$, $1 \le l \le J$. We then show that these approximations are sufficiently close to the corresponding numbers z and z' of class 1 and class $0'$ descendants of individuals A and B in generation r, so that $\mathbf{P}[S(z, z', w) = 0]$ is correspondingly close to $\exp\{-\lambda(\tilde{z}, \tilde{z}', w)\}$.

6 Approximating inter-point distances

We now return to the problem of real interest, the distribution of the graph distance $D := D_{k_1,k_2}$ between two vertices $(k_1, 1)$ and $(k_2, 1)$ in the intersection graph, taken to be infinite if the vertices are in different components.

6.1 Conditioning on the branching process

We begin by approximating the conditional probability $\mathbf{P}[D > d \,|\, Z]$ that A and B are more than distance d apart, given the trajectory of the bipartite process Z starting from A and B. The conditional probability is then a function only of the way in which the labels were assigned to the individuals in the process Z. The labelling determines the classes of the individuals, and the event $\{D > d\}$ occurs exactly when there are no overlaps between the labels of the class 1 and class $0'$ individuals that are descended from A and those of the descendants of B, at any generation l, $1 \le l \le d$, of Z. Let $\mathcal{G}_s$ denote the information in the labels up to generation s.

Proposition 6.1 *For any $1 \le l \le d$,*

$$\left|\mathbf{P}[D > l \,|\, Z, \mathcal{G}_{l-1} \cap \{D \ge l\}] - e^{-\lambda'(l,Z)}\right| \;\le\; \phi(l, Z), \tag{6.1}$$

where ϕ is given in (6.6) and (6.11), and $\lambda'(l,Z)$ in (6.8) and (6.10). It then follows in particular that

$$\left|\mathbf{P}[D > d \,|\, Z] - \prod_{l=1}^{d} e^{-\lambda'(l,Z)}\right| \;\le\; \sum_{l=1}^{d} \phi(l, Z). \tag{6.2}$$

Proof: Suppose, first, that l is even. By Proposition 5.1, the probability

$$\mathbf{P}[D > l \,|\, Z, \mathcal{G}_{l-1} \cap \{D \ge l\}]$$

is close to $\exp\{-\lambda(z(l), z'(l), \bar{n}, K)\}$, with $\bar{n}_k = n_k$, $z(l)$ the numbers of children of the different types of class 1 descendants of A in generation $l-1$, and $z'(l)$ is the same for descendants of B. When applying Proposition 5.1, R_l represents the number of class 1 descendants of A in generation $l-1$, and z_{lr} the number of offspring of the r-th of these; these offspring make up the class 1 and class $0'$ descendants of A in generation l. The error in the approximation is then no larger than

$$B_1(z(l), z'(l), \bar{n}, K) + B_1^*(z(l), z'(l), \bar{n}, \bar{n}^*(l-1), K),$$

a quantity that we shall need to bound later, where $\bar{n}_k^*(l-1)$ is the number of labels for $(k,1)$ individuals already used up to generation $l-1$ of the Z-process.

The quantities $z(l)$ and $z'(l)$ appearing in λ are not directly accessible, and are not functions of Z alone. However, we can exploit Proposition 5.2, provided that we can find suitable approximations to them. The first is to replace $z(l)$ by $X^A(l/2)$, the numbers of descendants of A of the different types in generation $l/2$ of the X-process, and $z'(l)$ by $X^B(l/2)$, noting that

$$0 \;\le\; X_k^A(l/2) - z_k(l) \;\le\; G_k^X(l/2); \quad 0 \;\le\; X_k^B(l/2) - z_k'(l) \;\le\; G_k^X(l/2), \tag{6.3}$$

and that the number of ghosts $G_k^X(l/2)$, investigated in Section 4, is calculated for the whole bivariate process. Then the quantities $X^A(l/2)$ and $X^B(l/2)$ can in turn be more simply approximated, using Theorem 3.6, by $\tau^{l/2}W^A\mu$ and $\tau^{l/2}W^B\mu$, where W^A is the limit of the martingale $\tau^{-i}\nu^T X^A(i)$, and W^B the limit of $\tau^{-i}\nu^T X^B(i)$: note that these two random variables are independent, by the branching property. From Proposition (5.2), replacing $z(l)$ by $\tau^{l/2}W^A\mu$ and $z'(l)$ by $\tau^{l/2}W^B\mu$, we incur a further error of at most $B_2(X^A(l/2), X^B(l/2), \varepsilon^A(l), \varepsilon^B(l), \bar{n}, K)$, where

$$\varepsilon_k^A(l) \;=\; G_k^X(l/2) + E(l/2, X^A); \qquad \varepsilon_k^B(l) \;=\; G_k^X(l/2) + E(l/2, X^B), \tag{6.4}$$

with $E(l/2, X^A) = X^A(l/2) - \tau^{l/2}W^A\mu$ and $E(l/2, X^A)$ defined analogously. By Theorem 3.6,

$$\mathbf{E}\{E(i, X)\} \;\le\; c_9((i+1)\gamma^i)^{1/2}, \tag{6.5}$$

for $X = X^A$ and for $X = X^B$.

We also clearly have $\bar{n}^*(l-1) \le T^X((l-2)/2)$, componentwise, where

$$T^X(s) \;:=\; \sum_{r=0}^{s} X(r) = \sum_{r=0}^{s}(X^A(r) + X^B(r)),$$

and, as observed above, $z(l) \le X^A(l/2)$, $z'(l) \le X^B(l/2)$. As a result, the approximation error at this step is no larger than

$$\begin{aligned}\phi(l, Z) \;:=\; & B_1(X^A(l/2), X^B(l/2), \bar{n}, K) + B_1^*(X^A(l/2), X^B(l/2), \bar{n}, T^X((l-2)/2), K) \\ & + B_2(X^A(l/2), X^B(l/2), \varepsilon^A(l), \varepsilon^B(l), \bar{n}, K),\end{aligned} \tag{6.6}$$

with $\varepsilon^A(l), \varepsilon^B(l)$ as in (6.4); thus we have, for l even,

$$|\mathbf{P}[D > l \,|\, Z, \mathcal{G}_{l-1} \cap \{D \ge l\}] - \exp\{-\lambda'(l, Z)\}| \;\le\; \phi(l, Z), \tag{6.7}$$

where, for l even,

$$\lambda'(l, Z) \;:=\; \lambda(\tau^{l/2}W^A\mu, \tau^{l/2}W^B\mu, \bar{n}, K). \tag{6.8}$$

A similar argument for l odd yields

$$|\mathbf{P}[D > l \,|\, Z, \mathcal{G}_{l-1} \cap \{D \ge l\}] - \exp\{-\lambda'(l, Z)\}| \;\le\; \phi(l, Z), \tag{6.9}$$

where, for l odd,

$$\lambda'(l, Z) \;:=\; \lambda(\zeta\tau^{(l-1)/2}W^A\tilde{\mu}, \zeta\tau^{(l-1)/2}W^B\tilde{\mu}, \bar{m}, J), \tag{6.10}$$

$$\begin{aligned}\phi(l, Z) \;:=\; & B_1(Y^A((l+1)/2), Y^B((l+1)/2), \bar{m}, J) \\ & + B_1^*(Y^A((l+1)/2), Y^B((l+1)/2), \bar{m}, T^Y((l-1)/2), J) \\ & + B_2(Y^A((l+1)/2), Y^B((l+1)/2), \eta^A(l), \eta^B(l), \bar{m}, J).\end{aligned} \tag{6.11}$$

Here,

$$\eta_k^A(l) \;=\; G_k^Y((l+1)/2) + E'((l+1)/2, Y^A); \qquad \eta_k^B(l) \;=\; G_k^Y((l+1)/2) + E'((l+1)/2, Y^B), \tag{6.12}$$

and

$$\mathbf{E}\{E'(i, Y)\} \;\le\; c_{14}u_{mn}(i\gamma^i)^{1/2},$$

for $Y = Y^A$ and for $Y = Y^B$, by Theorem 3.7. This proves the first statement of the proposition.

The second part is easier. We first note that

$$\mathbf{P}[D > l \mid Z] \;=\; \mathbf{E}\{\mathbf{E}(I[D > l-1]\mathbf{P}[D > l \mid Z, \mathcal{G}_{l-1} \cap \{D \ge l\}] \mid \mathcal{G}_{l-1}, Z) \mid Z\},$$

and deduce from the first part that

$$\bigl|\mathbf{P}[D > l \mid Z] - e^{-\lambda'(l,Z)}\mathbf{P}[D > l-1 \mid Z]\bigr| \;\le\; \phi(l,Z),$$

from which the last part follows. □

Thus, combining (6.7) and (6.9) with Proposition 6.1, we find that

$$|\mathbf{P}[D > d \mid Z] - \exp\{-W^A W^B L(d)\}| \;\le\; \sum_{l=1}^{d} \phi(l,Z), \tag{6.13}$$

where, using (3.3),

$$\begin{aligned}
L(2i) \;&:=\; \frac{\tau^{2i}-1}{\tau^2-1}\left\{\tau^2 \sum_{k=1}^{K} \frac{\mu_k^2}{n_k} + \zeta^2 \sum_{j=1}^{J} \frac{\tilde{\mu}_j^2}{m_j}\right\} \\
&=\; \frac{\tau^{2i}-1}{\tau^2-1}\,\tau(\tau+1) \sum_{k=1}^{K} \frac{\mu_k^2}{n_k} \;=\; \left(\frac{\tau}{\tau-1}\right) n^{-1}(\tau^{2i}-1) \sum_{k=1}^{K} \frac{\mu_k^2}{q_k^X}; \\
L(2i+1) \;&:=\; \frac{1}{\tau^2-1}\left\{(\tau^{2i+2}-\tau^2) \sum_{k=1}^{K} \frac{\mu_k^2}{n_k} + (\tau^{2i+2}-1)\zeta^2 \sum_{j=1}^{J} \frac{\tilde{\mu}_j^2}{m_j}\right\} \\
&=\; \frac{1}{\tau^2-1}\left\{(\tau^{2i+2}-\tau^2+\tau^{2i+3}-\tau) \sum_{k=1}^{K} \frac{\mu_k^2}{n_k}\right\} \\
&=\; \left(\frac{\tau}{\tau-1}\right) n^{-1}(\tau^{2i+1}-1) \sum_{k=1}^{K} \frac{\mu_k^2}{q_k^X},
\end{aligned}$$

and hence, with (2.3),

$$L(d) \;=\; \kappa n^{-1}(\tau^d - 1), \tag{6.14}$$

for d both even and odd.

6.2 The unconditional distribution

The unconditional probabilities for D are now given by taking expectations in conjunction with (6.13), so that it just remains to evaluate the terms $\mathbf{E}_{k_1,k_2}\phi(l,Z)$. Note that, in the approximation, randomness comes in only through the independent random variables W^A and W^B, the first with a distribution which depends only on the value of k_1, and the second on k_2.

To assist in judging the impact of the various factors in our bounds, it is convenient to define

$$i_0 \;:=\; \left\lfloor \frac{\log n}{\log \tau} \right\rfloor, \tag{6.15}$$

so that $\tau^{i_0} \le n < \tau^{i_0+1}$, and to set

$$\tau^{-1} \;<\; \varphi(n) \;:=\; n^{-1}\tau^{i_0} \;\le\; 1.$$

Then, with κ given in (2.3) and under standard asymptotics,

$$|L(d) - \tau^{d-i_0}\kappa\varphi(n)| \;\le\; \frac{\rho^X}{n}\left(\frac{\tau}{\tau-1}\right) \;\to\; 0, \tag{6.16}$$

and κ remains bounded away from 0 and ∞.

Theorem 6.2 *For $d = i_0 + u$, with $u \in \mathbb{Z}$ and $|u| < i_0/2$, we have*

$$\begin{aligned}|\mathbf{P}_{k_1,k_2}[D - i_0 > u] - \mathbf{E}_{k_1,k_2}\exp\{-W^A W^B \kappa\tau^u\varphi(n)\}|\\ \le\; c_{25}\bigl\{(\tau^{3u/2}+1)(n^{1/4}e(m,n)^2 \wedge 1) + (\tau^u+1)n^{1/4}e(m,n)\tilde{\theta}_{i_0}\bigr\},\end{aligned}$$

for a suitable quantity c_{25}, where

$$\tilde{\theta}_i \;:=\; (i+1)^{1/2}(\gamma/\tau^2)^{i/4}. \tag{6.17}$$

Proof: The approximating expression is immediate, from (6.13) and (6.16), incurring an error of at most

$$\nu_{k_1}\nu_{k_2}\frac{\rho^X}{n}\left(\frac{\tau}{\tau-1}\right).$$

For the rest, we just need to investigate $\mathbf{E}_{k_1,k_2}\phi(l,Z)$ for $1 \le l \le i_0 + u$.

To start with, for $l = 2r$, we have

$$\mathbf{E}_{k_1,k_2}B_1(X^A(r), X^B(r), \bar{n}, K) \;\le\; 4n^{-1}Kc_0\tau^r\rho^X \;\le\; n^{-1}c_{22}\tau^r, \tag{6.18}$$

with $c_{22} := 4Kc_0\rho^X$, from (5.1) and (3.11). Then

$$\begin{aligned}&\mathbf{E}_{k_1,k_2}B_1^*(X^A(r), X^B(r), \bar{n}, T^X(r-1), K)\\ &\quad= \sum_{k=1}^{K}\sum_{s=1}^{r-1}(nq_k^X)^{-2}\mathbf{E}_{k_1,k_2}\{X_k^A(r)X_k^B(r)[X_k^A(s) + X_k^B(s)]\}\\ &\quad\le\; c_{23}n^{-2}\tau^{3r},\end{aligned} \tag{6.19}$$

where $c_{23} = 2K(c_0\rho^X)^2(c_5\|\mu\|_\infty + c_6K\theta)/(\tau-1)$, from (3.9) and Corollary 3.3. For $\mathbf{E}_{k_1,k_2}B_2(X^A(r), X^B(r), \varepsilon^A(2r), \varepsilon^B(2r), \bar{n}, K)$, with ε^A and ε^B defined as in (6.4), we need to be a little more careful, because of the product $G_k^X(r)(X_k^A(r) + X_k^B(r))$. However, from Theorem 4.1, it follows by Markov's inequality that, for any $\Phi = \Phi(m,n,r)$,

$$\mathbf{P}_{k_1,k_2}\left[\max_{1\le k\le K} G_k^X(r) > \Phi\right] \;\le\; \mathbf{P}_{k_1,k_2}\left[\sum_{k=1}^{K} G_k^X(r) > \Phi\right] \;\le\; \Phi^{-1}\tau^{2r}Kc_{15}^*e(m,n)^4, \tag{6.20}$$

and because B_2 can never exceed the value 3, it follows that

$$\begin{aligned}&\mathbf{E}_{k_1,k_2}B_2(X^A(r), X^B(r), \varepsilon^A(2r), \varepsilon^B(2r), \bar{n}, K)\\ &\quad\le\; \mathbf{E}_{k_1,k_2}B_2(X^A(r), X^B(r), \tilde{\varepsilon}^A(2r), \tilde{\varepsilon}^B(2r), \bar{n}, K) + 6\mathbf{P}_{k_1,k_2}\left[\max_{1\le k\le K} G_k^X(r) > \Phi\right]\\ &\quad\le\; \mathbf{E}_{k_1,k_2}B_2(X^A(r), X^B(r), \tilde{\varepsilon}^A(2r), \tilde{\varepsilon}^B(2r), \bar{n}, K + 6c_{24}\Phi^{-1}\tau^{2r}e(m,n)^4,\end{aligned} \tag{6.21}$$

with

$$\tilde{\varepsilon}_k^A(2r) \;=\; \Phi + E(r, X^A) \quad \text{and} \quad \tilde{\varepsilon}_k^B(2r) \;=\; \Phi + E(r, X^B),$$

and with $c_{24} := K c_{15}^*$. But now, from (6.5) and (3.11), it follows that

$$\begin{aligned} &\mathbf{E}_{k_1,k_2} B_2(X^A(r), X^B(r), \tilde{\varepsilon}^A(2r), \tilde{\varepsilon}^B(2r), \bar{n}, K) \qquad (6.22) \\ &\quad\le\; 2K\rho^X n^{-1} c_0 \tau^r \{\Phi + c_9((r+1)\gamma^r)^{1/2}\} + 2n^{-1}\{\Phi^2 + c_9^2(r+1)\gamma^r\} \sum_{k=1}^{K} \frac{1}{q_k^X}. \end{aligned}$$

Choosing $\Phi^2(m,n,r) := \tau^r n e(m,n)^4$, and then adding the contributions from (6.18) – (6.22) for $1 \le r \le \lfloor (i_0+u)/2 \rfloor$ gives, after some computation, a bound of the form

$$c_{25}' \big(n^{1/4}(\tau^{3u/4}+1) e(m,n)^2 + (\tau^u + 1)\tilde{\theta}_{i_0} + (\tau^{3u/2}+1) n^{-1/2} \big). \qquad (6.23)$$

Bounds analogous to (6.18) and (6.19) hold also for $l = 2r-1$, with Y, J, m replacing X, K, n throughout the argument and estimates, and with c_{22} and c_{23} replaced by $c_{22}'\sqrt{\frac{m}{n}}$ and $c_{23}'\sqrt{\frac{m}{n}} u_{mn}^2$, where

$$c_{22}' \;=\; 4c_0 \mathfrak{z} \tau^{-1} J \rho^Y \quad \text{and} \quad c_{23}' \;=\; 2JZ(c_0\rho^Y/\tau)^2 (c_5 \mathfrak{z}^2 \tau^{-2} \|\tilde{\mu}\|_\infty + c_6' J\theta)/(\tau - 1).$$

The bound corresponding to (6.20) is

$$\mathbf{P}_{k_1,k_2}\left[\max_{1\le j\le J} G_j^Y(r) > \Phi \right] \;\le\; \Phi^{-1}\sqrt{\frac{m}{n}}\, \tau^{2(r-1)} J c_{16}^* e(m,n)^4\,, \qquad (6.24)$$

and we also have

$$\begin{aligned} &\mathbf{E}_{k_1,k_2} B_2(Y^A(r), Y^B(r), \tilde{\eta}^A(2r-1), \tilde{\eta}^B(2r-1), \bar{m}, J) \qquad (6.25) \\ &\quad\le\; 2J\rho^Y m^{-1}\tau^{-1} c_0 \tau^r \mathfrak{z} \sqrt{\frac{m}{n}}\, \{\Phi + c_{14} u_{mn}(r\gamma^r)^{1/2}\} + 2m^{-1}\{\Phi^2 + (c_{14}u_{mn})^2 r\gamma^r\} \sum_{j=1}^{J} \frac{1}{q_j^Y}, \end{aligned}$$

with

$$\tilde{\eta}_k^A(2r-1) \;=\; \Phi + E'(r, Y^A) \quad \text{and} \quad \tilde{\eta}_k^B(2r-1) \;=\; \Phi + E'(r, Y^B).$$

Here, we take $\Phi^2 := m\tau^r e(m,n)^4$ in (6.20) and (6.25), and then, adding the errors over $1 \le r \le \lfloor i_0 + 1 + u \rfloor$, and after much calculation, a bound of the form

$$c_{25}'' \left\{ n^{1/4} e(m,n)^2 (\tau^{3u/4}+1) + n^{1/4} e(m,n)(\tau^u+1)\tilde{\theta}_{i_0} + (\tau^{3u/2}+1)\sqrt{\frac{n}{m}} e(m,n)^2 \right\} \qquad (6.26)$$

is obtained.

To deduce the bound given in the theorem, it now suffices to observe that

$$\big(n^{1/4} e(m,n)^2 \big)^2 \;\ge\; n^{-1/2} + \sqrt{\frac{n}{m}}\, e(m,n)^2,$$

so that the final terms in (6.23) and (6.26) can be absorbed into the first term, if the larger of the τ-exponents is used. □

The defective real valued random variable U, whose distribution function

$$\mathbf{P}_{k_1,k_2}[U \le u] \;=\; 1 - \mathbf{E}_{k_1,k_2} \exp\{-W^A W^B \kappa \tau^u \varphi(n)\} \tag{6.27}$$

approximates that of $D - i_0$ for integer arguments, can be expressed as a (defective) translation mixture of scaled negative standard Gumbel random variables. If $W^A W^B$ has distribution function F_{k_1,k_2} on $\mathbf{R}_+$, and if

$$\mathbf{P}_{k_1,k_2}[U' \le u] \;:=\; \int_{(0,\infty)} \mathbf{P}[-(\log\tau)^{-1}(\Gamma + \log x + \log\kappa) \le u]\, dF_{k_1,k_2}(x), \tag{6.28}$$

where Γ denotes a standard Gumbel random variable, then

$$\mathbf{P}_{k_1,k_2}[U \le u] \;=\; \mathbf{P}_{k_1,k_2}[U' \le u + \log\varphi(n)/\log\tau].$$

Alternatively, we can write

$$\mathbf{P}[U = \infty] \;=\; \mathbf{P}[U' = \infty] \;=\; F_{k_1,k_2}(0) \;=\; 1 - \mathbf{P}_{k_1}[W > 0]\mathbf{P}_{k_2}[W > 0], \tag{6.29}$$

and express the distribution $\mathcal{L}(U' \,|\, U' < \infty)$ as that of a random variable $\widetilde{U}$, realized as

$$\widetilde{U} \;=\; -\frac{1}{\log\tau}\{\Gamma + \log\widetilde{W}_A + \log\widetilde{W}_B + \log\kappa\}, \tag{6.30}$$

where Γ, $\widetilde{W}_A$ and $\widetilde{W}_B$ are independent,

$$\mathbf{P}[\widetilde{W}_A \le w] \;=\; \mathbf{P}_{k_1}[W \le w \,|\, W > 0] \quad\text{and}\quad \mathbf{P}[\widetilde{W}_B \le w] \;=\; \mathbf{P}_{k_2}[W \le w \,|\, W > 0].$$

Note that $F_{k_1,k_2}(0) = \mathbf{P}[U' = \infty]$ indeed approximates the probability that $(k_1, 1)$ and $(k_2, 1)$ are in different components of the graph, and are hence at infinite distance from one another, as can be seen in the following result.

Theorem 6.3 *There are constants $\tau_1 > 1$ and $c_{26} < \infty$ such that*

$$\max\{\mathbf{P}_{k_1,k_2}[D < \infty \,|\, W^A = 0], \mathbf{P}_{k_1,k_2}[D < \infty \,|\, W^B = 0]\} \le c_{26}\tau_1^{-i_0}.$$

PROOF: We make the calculation for A; for B the argument is the same. From the general theory of multi-type branching processes, see for example [7] or [12], conditional on the event $\{W^A = 0\}$, X^A is a subcritical branching process, and there exist $\tau_1 > 1$ and $C < \infty$ such that

$$\mathbf{E}_k[\|X^A(i)\|_1 \,|\, W^A = 0] \;\le\; C\tau_1^{-i} \quad \text{for all } k,$$

and thus

$$\mathbf{E}_k[\|Y^A(i)\|_1 \,|\, W^A = 0] \;\le\; C\tau_1^{-i+1} c_0 \zeta \quad \text{for all } k.$$

Hence, immediately,

$$\mathbf{P}_{k_1,k_2}[2i_0 < D < \infty \,|\, W^A = 0] \;\le\; C\tau_1^{-i_0}.$$

Then, for $1 \leq i \leq i_0$, from (3.11),

$$\mathbf{P}_{k_1,k_2}[D = 2i \,|\, W^A = 0] \;\leq\; \sum_{k=1}^{K} n_k^{-1} \mathbf{E}_{k_1,k_2} X_k^A(i) \mathbf{E}_{k_1,k_2} X_k^B(i) \;\leq\; n^{-1} K C \tau_1^{-i} c_0 \tau^i \rho^X;$$
$$\mathbf{P}_{k_1,k_2}[D = 2i-1 \,|\, W^A = 0] \;\leq\; m^{-1} J C \zeta^2 \tau_1^{-i+1} c_0 \tau^{i-1} \rho^Y \;\leq\; n^{-1} J C \mathfrak{z}^2 \tau_1^{-i+1} c_0 \tau^{i-1} \rho^Y,$$

and the theorem follows by adding over $1 \leq i \leq i_0$. □

In view of the considerations above, our approximation can be summarized as follows.

Corollary 6.4 *For $d = i_0 + u$, with $u \in \mathbb{Z}$ and $|u| < i_0/2$, we have*

$$|\mathbf{P}_{k_1,k_2}[D \leq u + i_0] - \mathbf{P}_{k_1,k_2}[U' \leq u + \log \varphi(n)/\log \tau]| \;\leq\; \delta(\tau^u, m, n);$$
$$|\mathbf{P}_{k_1,k_2}[D = \infty] - \mathbf{P}_{k_1,k_2}[U' = \infty]|$$
$$\leq\; \delta(n^\alpha, m, n) + n^{-1} + \mathbf{P}_{k_1,k_2}[0 < W^A W^B \leq \tau^2 \kappa^{-1} n^{-\alpha} \log n] + 2c_{26} \tau_1^{-i_0},$$

for any $0 < \alpha < (i_0 - 2)/2i_0 \approx 1/2$, where

$$\delta(y, m, n) \;:=\; c_{25} \big\{ (y^{3/2} + 1)(n^{1/4} e(m,n)^2 \wedge 1) + (y+1) n^{1/4} e(m,n) \tilde{\theta}_{i_0} \big\},$$

i_0 is as in (6.15), $\tilde{\theta}_i$ is as in (6.17), τ_1 is as for Theorem 6.3 and U' has distribution given either by (6.28) or by (6.29) and (6.30).

Proof: The first inequality is from Theorem 6.2. For the second, we have

$$\mathbf{P}_{k_1,k_2}[D < \infty]$$
$$\leq\; \mathbf{P}_{k_1,k_2}[D < \infty \,|\, W^A = 0] + \mathbf{P}_{k_1,k_2}[D < \infty \,|\, W^B = 0] + \mathbf{P}_{k_1,k_2}[W^A W^B > 0],$$

giving

$$\mathbf{P}_{k_1,k_2}[D = \infty] \;\geq\; 1 - \mathbf{P}_{k_1,k_2}[W^A W^B > 0] - 2c_{26} \tau_1^{-i_0} \;=\; \mathbf{P}_{k_1,k_2}[U' = \infty] - 2c_{26} \tau_1^{-i_0}.$$

On the other hand, taking $u = \lfloor \alpha i_0 \rfloor$ in the first part, we have

$$\mathbf{P}_{k_1,k_2}[D = \infty] \;\leq\; 1 - \mathbf{P}_{k_1,k_2}[U \leq u] + \delta(n^\alpha, m, n),$$

and, from (6.27), for any $C > 0$,

$$\mathbf{P}_{k_1,k_2}[U \leq u] \;\geq\; 1 - \mathbf{P}_{k_1,k_2}[W^A W^B \leq C n^{-\alpha} \log n] - \exp\{-C n^{-\alpha} \log n \kappa (n^\alpha/\tau) \varphi(n)\}$$
$$\geq\; \mathbf{P}_{k_1,k_2}[W^A W^B > 0] - \mathbf{P}_{k_1,k_2}[0 < W^A W^B \leq C n^{-\alpha} \log n]$$
$$- \exp\{-C \log n \kappa/\tau^2\}.$$

Hence, taking $C = \tau^2/\kappa$,

$$\mathbf{P}_{k_1,k_2}[D = \infty]$$
$$\leq\; \mathbf{P}_{k_1,k_2}[U' = \infty] + \mathbf{P}_{k_1,k_2}[0 < W^A W^B \leq \tau^2 \kappa^{-1} n^{-\alpha} \log n] + n^{-1} + \delta(n^\alpha, m, n),$$

and the corollary is proved. □

Remark. The corresponding result for the unipartite Erdős–Rényi graph may also be of interest, although, as discussed at the end of Section 1, it is not directly useful for our purposes. For such a graph, the vertices are divided into K types, with n_k of type k, $1 \leq k \leq K$, and with $n := \sum_{k=1}^K n_k$. Edges are then independently assigned, with probabilities $p_{k,k'}$ depending on the vertex types k and k': the matrix P is thus symmetric. The mean matrix M for the associated branching process is given by PN, where $N := \mathrm{diag}\{n_1, \ldots, n_K\}$, and we assume that it is irreducible and aperiodic, and that its largest eigenvalue $\tilde{\tau} > 1$. With these assumptions, and writing μ^T for the left eigenvector of M with eigenvalue $\tilde{\tau}$, only small changes need to be made to the sketched argument concluding Section 2. Considering coincidences in the indices in order of increasing generation number, and with the offspring of A considered before those of B, links in the Erdős–Reńyi graph arise exactly when there are coincidences between indices of the $X_k^A(i)$ and those of the $X_k^B(i-1)$, or between indices of the $X_k^B(i)$ and those of the $X_k^A(i)$. This leads to an approximate mean number of coincidences, up to and including the time when the first i generations of descendants of A and the first $i-1$ of B have been considered, of $\tilde{\kappa} n^{-1} \tilde{\tau}^{2i-1} W^A W^B$, where

$$\tilde{\kappa} \ := \ \frac{\tilde{\tau}}{\tilde{\tau} - 1} \sum_{k=1}^K \frac{\mu_k^2}{q_k},$$

and $q_k := n_k/n$. For the time until the first i generations of both have been considered, the corresponding approximation is $\tilde{\kappa} n^{-1} \tilde{\tau}^{2i} W^A W^B$. This gives the probability that the distance between A and B exceeds d as being approximately

$$\mathbf{E}_{k_1,k_2} \left\{ e^{-\tilde{\kappa} n^{-1} \tilde{\tau}^d W^A W^B} \right\}, \tag{6.31}$$

very much the same as the formula in Theorem 6.2. Note once again that the assumption of irreducibility prevents this line of argument being directly applicable to the bipartite model.

6.3 Asymptotic behaviour

Recalling the standard asymptotics of Section 3, we now distinguish the possibilities for the bipartite branching process starting with a single vertex of type $(k, 1)$, according to the behaviour of the ratio m/n, as $n \to \infty$. This, in turn, enables one to deduce the asymptotic form of the approximating random variable U'.

First, note that the $Y_j^{(m,n)}(1) \sim \mathrm{Bi}\,(mq_j^Y, (mn)^{-1/2}\Pi_{kj})$, $1 \leq j \leq J$, are independent. If $m/n \to r$ with $0 < r < \infty$, then the Poisson approximation to the binomial distribution thus shows that the distribution of $Y_j^{(m,n)}(1)$ differs in total variation from $\mathrm{Po}\,(\sqrt{r}\Pi_{kj} q_j^Y(m,n))$ by at most $(mn)^{-1/2}\Pi_{kj} \sim n^{-1}\Pi_{kj}/\sqrt{r}$, and, conditional on $Y^{(m,n)}(1)$, the distributions of the $X_l^{(m,n)}(1)$ are independent, and close to the same order to $\mathrm{Po}\,(\sum_{j=1}^J Y_j^{(m,n)}(1)\Pi_{lj} q_l^X(m,n) r^{-1/2})$. Hence, as m and n tend to infinity in this way, the bipartite branching process converges to the one with exactly Poisson offspring distributions and with $Q_X^{(m,n)}$ and $Q_Y^{(m,n)}$ replaced by Q_X and Q_Y. Hence the distribution $\mathcal{L}(W^{(m,n)} \,|\, X(0) = e^{(k)})$ converges to $\mathcal{L}(W \,|\, X(0) = e^{(k)})$, for each k, where W is the limiting random variable associated with the limiting Poisson–based branching process. It thus follows that the distribution of the random variable $U'^{(m,n)}$ also converges to that of the corresponding U'. However, the distribution of $U^{(m,n)}$ does not converge in general, because the value of $\log\varphi(n)/\log\tau$ oscillates between -1 and 0 as n varies.

If $m/n \to \infty$, the Poisson approximation $\mathrm{Po}\,(\sqrt{m/n}\Pi_{kj}q_j^Y(m,n))$ to the distribution of $Y_j^{(m,n)}(1)$ still has error of at most $(mn)^{-1/2}\Pi_{kj}$. However, a simple calculation shows that, for $B_j^{(m,n)}$ a Bernoulli random vector with $\mathbf{P}[B_j^{(m,n)} = e^{(l)}] = q_l^X(m,n)\Pi_{lj}\sqrt{n/m}$,

$$d_{TV}\big(\mathcal{L}(X^{(m,n)}(1) \,|\, Y^{(m,n)}(1) = \tilde{e}^{(j)}), \mathcal{L}(B_j^{(m,n)})\big) \;\le\; \left\{\sum_{k=1}^K q_k^X(m,n)\Pi_{kj}\sqrt{n/m}\right\}^2 .$$

It now follows from the Poisson thinning theorem (see for example Chapter 8 Section 6 in [9]), and because

$$\sum_{j=1}^J \Pi_{kj}q_j^Y(m,n)\Pi_{lj}q_l^X(m,n) \;=\; M_X^{(m,n)}(k,l),$$

that

$$\begin{aligned} &d_{TV}\Big(\mathcal{L}(X^{(m,n)}(1) \,|\, X^{(m,n)}(0) = e^{(k)}), \otimes_{l=1}^K \mathrm{Po}\,(M_X^{(m,n)}(k,l))\Big) \\ &\quad\le\; \sum_{j=1}^J \left(\frac{\Pi_{kj}}{\sqrt{mn}} + \sqrt{\frac{n}{m}}\Pi_{kj}q_j^Y(m,n)\left\{\sum_{k=1}^K q_k^X(m,n)\Pi_{kj}\right\}^2\right), \end{aligned}$$

an error of order $O(\sqrt{n/m})$. Thus, in this regime, the offspring distribution for the $X^{(m,n)}$ process, which determines the distribution of $W^{(m,n)}$, approaches one with independent Poisson components, having means given by the matrix M_X. Again, this entails the convergence of $U'^{(m,n)}$ to U', but not the convergence of $U^{(m,n)}$.

Finally, if m/n is small, the simple bound $(1-p)^l \ge 1 - lp$ shows that

$$\mathbf{P}[Y_j^{(m,n)}(1) \ne 0] \le (m/n)^{1/2}\Pi_{kj}q_j^Y(m,n),$$

from which it follows that $\mathbf{P}[W > 0] \le (m/n)^{1/2}\sum_{j=1}^J q_j^Y(m,n)\Pi_{kj}$. Hence, for $m/n \to 0$, the distance between two randomly chosen vertices $(k_1, 1)$ and $(k_2, 1)$ is infinite, with probability close to 1. However, if the two vertices A and B do each have an edge joining them to the object set, then each is connected to just one object with conditional probability of order $1 - O(\sqrt{\frac{m}{n}})$, and the objects to which they are linked are distinct with probability of order $1 - O(1/m)$. The distance between these two objects can now be investigated, in this regime, by swapping the roles of vertices and objects, and using the theorems above.

Thus if, in this scheme, m/n converges to a finite or infinite limit, the approximating probability distributions $\mathcal{L}(U'^{(m,n)})$ remain relatively stable. In the error terms, the quantities $\tau^{(m,n)}$ and $\gamma^{(m,n)}$ converge to limits τ and γ, the corresponding quantities for the limit matrix M_X. The factor $n^{1/4}e(m,n)^2$ behaves like $n^{-1/4}$, and $\tilde{\theta}_{i_0}$ like $n^{-\delta}\log n$, for some δ depending on M_X, as long as m/n is bounded below as $n \to \infty$; in view of (3.6) and (6.17), it follows that $\delta \leq 1/4$. The discussion above shows that m/n bounded below is the case of main interest.

7 An exponential random graph model

Rank 1 matrices $P = \alpha\beta^T$ give rise to an exponential random graph model. The individual edges are independent, as before, and the probability of a vertex of type i connecting to an object of type j is of product form. These models have been extensively studied in the social science literature, see for example [20] and references therein; for applications to affiliation networks as bipartite networks, see for example [22].

In this case,

$$M_X(k,l) \;=\; \sum_{j=1}^{J} \alpha_k \beta_j m_j \beta_j \alpha_l n_l \;=\; \alpha_k \alpha_l n_l \sum_{j=1}^{J} m_j \beta_j^2,$$

so that

$$M_X \;=\; C\alpha\alpha^T N_X,$$

with $C = \sum_{j=1}^{J} m_j \beta_j^2$, has $\tau = C\alpha^T N_X \alpha$, $\mu = N_X\alpha / \mathbf{1}^T N_X \alpha$ and $\nu = C(\tau^{-1}\mathbf{1}^T N_X \alpha)\alpha$. Here, $\mathbf{1}$ is a $K \times 1$-vector of 1's.

As shown in the preceding sections, the principal eigenvalue τ of M_X in our general multitype intersection graph is of critical importance in determining network distances. It turns out that its value can be bounded below by that obtained from an associated rank 1 matrix, adding to the importance of the exponential models. To see this, set $D_k^X := \sum_{j=1}^{J} p_{kj} m_j$ to be the average degree of a type $(k,1)$ vertex, and write $s_B^2 := \sum_{k=1}^{K} n_k (D_k^X)^2$.

Proposition 7.1 *If the values D_k^X, $1 \leq k \leq K$, are fixed, then*

$$\tau \;\geq\; s_B^2/m,$$

and this value of τ is attained by taking $p_{kj} = D_k^X/m$ for all j; with this choice of the p_{kj}'s, a vertex makes no distinction as to which types of object it has links to. The lower bound is minimized, if $\sum_{k=1}^{K} D_k^X = D_+^X$ is fixed, by taking $D_k^X = D_+^X/n$, so that all links have the same probability $p_{kj} = D_+^X/(mn)$.

PROOF: The proof is taken from [2], p.15. The matrix $M_X := PN_Y P^T N_X$ has the same eigenvalues as the symmetric matrix

$$V \;:=\; N_X^{1/2} P N_Y P^T N_X^{1/2}.$$

Write $u_{kj} := (D_k^X)^{-1} m p_{kj} - 1$, and note that $\sum_{j=1}^J m_j u_{kj} = 0$ for each k. Then

$$\begin{aligned} V_{ki} &= \sqrt{n_k} \sum_{j=1}^J p_{kj} m_j p_{ij} \sqrt{n_i} \\ &= m^{-1} D_k^X \sqrt{n_k} D_i^X \sqrt{n_i} \left\{ 1 + m^{-1} \sum_{j=1}^J m_j u_{kj} u_{ij} \right\}. \end{aligned}$$

From the Rayleigh-Ritz Theorem [11], Theorem 4.2.2, it follows that τ, the largest eigenvalue of V, is at least as large as $e^T V e/(e^T e)$, for any $e \in \mathbf{R}^K$. Taking $e_k = D_k^X \sqrt{n_k}/s_B$, $1 \le k \le K$, gives

$$\tau \;\ge\; m^{-1} s_B^2 + \|v\|_2^2,$$

where

$$v_j \;:=\; m^{-1} \sum_{k=1}^K n_k (D_k^X)^2 u_{kj} \sqrt{m_j}, \qquad 1 \le j \le J.$$

Since, with $p_{kj} = D_k^X/m$ for all k, V takes the form $m^{-1} w w^T$, with $w = N_X^{1/2} D^X$, and hence has largest eigenvalue $m^{-1}\|w\|_2^2 = s_B^2/m$, this proves the first statement of the proposition; the second is now immediate. $\square$

Thus the value of τ for a given P is always bigger than that corresponding to as homogeneous a choice of the link probabilities as is allowed by the constraints on the average number of objects linked to a given vertex.

Any rank one choice $P = \alpha\beta^T$ has a minimality property, analogous to that of Proposition 7.1, but of a less intuitive nature. The matrix $P = \alpha\beta^T$ minimizes the maximum eigenvalue of M_X among all choices of P satisfying the constraint

$$P N_Y \beta \;=\; (\beta^T N_Y \beta)\alpha.$$

References

[1] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes.* Springer, New York. Math. Review 0373040

[2] Barbour, A.D. (1978). Macdonald's model and the transmission of bilharzia. *Trans. Roy. Soc. Trop. Med. Hyg.* **72**, 6–15.

[3] Barbour, A.D. and Eagleson, G.K. (1983). Poisson approximation for some statistics based on exchangeable trials. *Adv. Appl. Probab.* **15**, 585–600. Math. Review 0706618

[4] Barbour, A.D., Holst, L. and Janson, S. (1992). *Poisson Approximation.* Oxford University Press. Math. Review 1163825

[5] Barbour, A.D. and Reinert, G. (2001). Small Worlds. *Random Structures and Algorithms* **19**, 54–74. Math. Review 1848027. Correction: *ibid* **25**, 115 (2004). Math. Review 1848027

[6] Britton, T., Deijfen, M., Lageras, A., Lindholm, M. (2008). Epidemics on random graphs with tunable clustering. *J. Appl. Probab.* **45**, 743–756. Math. Review 2455182

[7] Daly, F. (1979). Collapsing supercritical branching processes. *Journal of Applied Probability* **16**, 732–739. Math. Review 0549553

[8] Daudin, J-J., Picard, F. and Robin, S. (2008). A mixture model for random graphs. *Statistics and Computing* **18**, 173–183. Math. Review 2390817

[9] Gut, A. (2009). *An Intermediate Course in Probability.* 2nd ed. Springer-Verlag, New York. Math. Review 2528081

[10] Harris, T.E. (1989). *The Theory of Branching Processes.* Dover, New York. Math. Review 1991122

[11] Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis.* Cambridge University Press. Math. Review 0832183

[12] Jagers, P. and Lageras, A.N. (2008). General branching processes conditioned on extinction are still branching processes. *Elect. Comm. in Probab.* **13**, 540–547. Math. Review 2453547

[13] Kesten, H. and Stigum, B.P. (1966). Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Ann. Math. Statist.* **37**, 1463–1481. Math. Review 0200979

[14] Kurtz, T., Lyons, R., Pemantle, R., Peres, Y. (1997). A Conceptual Proof Of The Kesten-Stigum Theorem For Multi-Type Branching Processes. In *Classical and Modern Branching Processes*, K. Athreya and P. Jagers (editors), Springer, New York, 181–186. Math. Review 1601737

[15] Mode, C. (1971). *Multitype Branching Processes: Theory and Applications.* Elsevier, New York. Math. Review 0279901

[16] Nowicki, K. and Snijders, T. (2001). Estimation and Prediction for Stochastic Blockstructures. *Journal of the American Statistical Association* **96**, 1077–1087. Math. Review 1947255

[17] Rasch, G. (1961). On general laws and the meaning of measurement in psychology. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, IV.* Berkeley, 321–333.

[18] Reinert, G. and Waterman, M.S. (2007). On the length of an exact position match in a random sequence. *Transactions on Computational Biology and Bioinformatics* **4**, 153–156.

[19] Robins, G. and Alexander, M. (2004). Small worlds among interlocking directors: network structure and distance in bipartite graphs. *Computational & Mathematical Organization Theory* **10**, 69–94.

[20] Robins, G., Snijders, T., Wang, P., Handcock, M., and Pattison, P. (2008). Recent developments in exponential random graph (p^*) models for social networks. *Statistics and Computing* **18**, 173–183.

[21] Tanay, A., Sharan, R., and Shamir, R. (2002). Discovering statistically significant biclusters in gene expression data. *Bioinformatics* **18**, (Suppl 1): S136–S144.

[22] Wang, P., Sharpe, K., Robins, G., and Pattison, P. (2009). Exponential random graph (p^*) models for affiliation networks. *Social Networks* **31**, 12–25.